

Knolle, Niels [Hrsg.]

Evaluationsforschung in der Musikpädagogik

formal überarbeitete Version der Originalveröffentlichung:

formally revised edition of the original source:

Essen : Die Blaue Eule 2010, 324 S. - (Musikpädagogische Forschung; 31)



Bitte verwenden Sie in der Quellenangabe folgende URN oder DOI /

Please use the following URN or DOI for reference:

urn:nbn:de:0111-pedocs-120399

10.25656/01:12039

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-120399>

<https://doi.org/10.25656/01:12039>

in Kooperation mit / in cooperation with:



<http://www.ampf.info>

Nutzungsbedingungen

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

We grant a non-exclusive, non-transferable, individual and limited right to using this document.

This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Kontakt / Contact:

peDOCS
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation
Informationszentrum (IZ) Bildung
E-Mail: pedocs@dipf.de
Internet: www.pedocs.de

Mitglied der


Leibniz-Gemeinschaft

**Musikpädagogische
Forschung**

**Niels Knolle
(Hrsg.)**

**Evaluationsforschung
in der Musikpädagogik**



Themenstellung: Evaluationsforschung ist zu einem bedeutsamen Zweig der Bildungsforschung geworden, die Vielfalt der Beiträge zur 31. AMPF-Tagung >Evaluationsforschung in der Musikpädagogik< macht deutlich, dass die musikpädagogische Forschung hierzu einen bedeutsamen Beitrag zu liefern in der Lage ist. So zielen die Beiträge dieses Bands darauf, die Voraussetzungen, Inhalte, Methoden und Resultate von musikunterrichtlichen Reformansätzen und Innovationen im Blick auf die mit ihnen verbundenen Ziele zu überprüfen und zu bewerten, um so zu einer Verbesserung des musikbezogenen Handelns bzw. entsprechender Lehr-Lern-Prozesse zu gelangen.

Der Herausgeber: *Niels Knolle*, geb. 1944. Arbeitsschwerpunkte: Multimedia als Instrument, Werkzeug und Thema des Musikunterrichts; Didaktik der Populären Musik; Bildungsreformen in der Musikpädagogik. Langjährige Arbeit in den Vorständen der BFG Musikpädagogik, des AMPF, der Bundesfachausschüsse >Musikpädagogik< und >Musik und Medien< des Deutschen Musikrats. 1999 - 2003 Mitherausgeber der Zeitschrift >Musik in der Schule<. Von 1996 bis 2010 Universitätsprofessor für Musikpädagogik an der Otto-von-Guericke-Universität Magdeburg.

Inhalt

Niels Knolle:

Vorwort 7

Beiträge zum Tagungsthema

Udo Kelle, Brigitte Metje:

Mixed Methods in der Evaluationsforschung. Das Verhältnis zwischen Qualität und Quantität in der Wirkungsanalyse 9

Susanne Naacke:

Schulentwicklung mit Chor- und Bläserklassen. Eine qualitative Fallstudie am „Evangelischen Gymnasium am Dom zu Brandenburg“ 41

Forschungspreis 2009 Hösbach

Jens Knigge, Anne Niessen, Anne-Katrin Jordan:

Erfassung der Kompetenz „Musik wahrnehmen und kontextualisieren“ mit Hilfe von Testaufgaben - Aufgabenentwicklung und -analyse im Projekt KoMus 81

Anne-Katrin Jordan, Andreas C. Lehmann, Jens Knigge:

Kompetenzmodellierung mit Methoden der Item-Response-Theorie (IRT) - Erste Ergebnisse der Validierung eines Modells für den Bereich „Musik wahrnehmen und kontextualisieren“ 109

Jürgen Oberschmidt:

Metaphorischer Sprachgebrauch im Unterricht - Überlegungen zur Evaluierung der Schülersprache 131

Kai Stefan Lothwesen:

Musikalisches Erleben und Lernen zwischen Musikschule und Grundschule. Methodenkritische Reflexionen am Beispiel der Evaluation des Programms „Monheimer Modell – Musikschule für alle“ 155

Dirk Bechtel:

„Wie Lehrer lieber lernen“ - Eine qualitative Studie über die Rolle von Fortbildungen aus der Sicht von Musiklehrerinnen und -lehrern 179

Eva Mödinger, Gabriele Hofmann:

Lampenfieber und Aufführungssängste bei Kindern und Jugendlichen - Erhebungen zur Selbstwahrnehmung im Rahmen musikalischer Vortragssituationen 201

Matthias Stubenvoll:

Qualität entsteht beim Lernen - Lerner integrierende Qualitätsbeurteilung beim E-Learning 211

Wibke Gütay:

Darf es noch ein bisschen mehr sein? Auswirkungen von Stimmtraining bei Chorklassenkindern 229

Freie Beiträge

Robert Lang:

Musiktheorie in musizierpraktischem Schulunterricht. Zur Effizienz basaler Harmonielehre für das Improvisieren mit Keyboards 255

Konsortium des JeKi-Forschungsschwerpunkts:

Der BMBF-Forschungsschwerpunkt zu „Jedem Kind ein Instrument“ in Nordrhein-Westfalen und Hamburg 275

Richard von Georgi, Kai Stefan Lothwesen:

Handlungskompetenzen und Studiumsmotivation von Musikstudierenden 305

Vorwort

Unter dem Eindruck der Ergebnisse der Bildungsstudien der vergangenen zehn Jahre – seien sie bezogen auf die Ermittlung von Kompetenzen von SchülerInnen in zentralen Schulfächern wie Mathematik, Deutsch oder Naturwissenschaften oder den Vergleich von Strukturen der Ausbildungs- bzw. Qualifizierungsprozesse in Kindertagesstätten, Schulen und Hochschulen – sind seitens der Öffentlichen Hand teilweise erhebliche Mittel investiert worden in Maßnahmen, die der Verbesserung der Situation im Bildungssektor dienen. Solche Maßnahmen setzen – idealtypisch gesehen – eine hinreichend ideologiefreie Situations- und Ursachenanalyse voraus und sie bedürfen einer längerfristig orientierten, umfassenden Reflektion ihrer bildungspolitischen Zielsetzungen und Chancen der Verwirklichung im föderalistischen Dickicht der Eigeninteressen der Bundesländer in Deutschland. Und, nicht minder wichtig und anspruchsvoll, einer unabhängigen, d.h. wissenschaftlichen Überprüfung ihrer Wirksamkeit unter dem Aspekt der Nachhaltigkeit.

Evaluationsforschung ist damit zu einem bedeutsamen Zweig der Bildungsforschung geworden. Im Kern zielt sie darauf, die Voraussetzungen, Inhalte, Methoden und Resultate von Reformen und Innovationen im Blick auf die mit ihnen verbundenen Ziele zu überprüfen und zu bewerten, um so zu einer Verbesserung zu gelangen. Reformen wie auch Innovationen ziehen, insofern als sie auf sehr unterschiedlichen inhaltlichen Ebenen angesiedelt sind, entsprechend unterschiedliche Fragestellungen und Strategien der Evaluation nach sich.

Konkret auf die beiden zentralen Arbeitsfelder der Musikpädagogik bezogen, also den produktiven und rezeptiven Umgang mit Musik und dessen Einfluss auf die Entwicklung mentaler Selbstbilder und musikbezogener Interaktionen sowie die Vermittlung von musikbezogenen Kompetenzen und instrumentalpraktischen Fähigkeiten im Blick auf die Nachhaltigkeit von Lehr-Lern-Prozessen, wird deutlich, dass hier ein weit gefasstes Verständnis von Evaluationsforschung in der Musikpädagogik unumgänglich ist (und insofern eine forcierte Öffnung hin zur Kooperation mit den Humanwissenschaften).

Die Beiträge zur 31. Jahrestagung des AMPF im Herbst 2009 spiegeln in ihrer thematischen und methodischen Vielfalt in bemerkenswerter Weise dieses Verständnis von Evaluationsforschung wider, ob es nun um die Reflexion und Bewertung von speziellen Musikschulprogrammen geht oder um die Er-

fahrungspotenziale der Schülersprache im Unterricht, oder um künstlerische Transformationsprozesse in der Lehre, die Evaluation eines Erklärungsmodells zur Schulentwicklung, oder um die Lern-Wirksamkeit von Aufgabenstellungen im Musikunterricht und nicht zuletzt um die Auswirkungen des Stimmtrainings bei Chorklassenkindern usw..

All diese Forschungsvorhaben lassen sich als begleitende Forschung zu musikpädagogischen Modellen, Verfahrensansätzen oder auch Innovationen verstehen. Sie sind insofern *herrschaftsfrei*, als sie sich nicht mit ihren Analysen und Bewertungen zu einem hierarchisch verwertbaren Werkzeug der Durchsetzung forschungsfremder Interessen eignen. Zu einem ganz anderen Bild gelangt man indes, wenn man den Blick auf Evaluationsprozesse wirft, die - bildungspolitisch motiviert - vom Staat initiiert sind (konkret von der Kultusministerkonferenz) und von privaten Firmen, z.B. Akkreditierungsagenturen, oder von Sondereinrichtungen der Kultusministerien durchgeführt werden. Man wird die in einigen Bundesländern seit einiger Zeit durchgeführten sog. Inspektionen der staatlichen Schulen wie auch die Akkreditierungen von Studiengängen an den Hochschulen nicht als Evaluationsforschung bezeichnen können, sie evaluieren aber Planungs- und Implementationsprozesse aus dem politischen Interesse heraus, Bildung als ‚Humankapital‘ - wie Richard Münch kürzlich gezeigt hat - für die Zwecke der ökonomischen Konkurrenz zu optimieren. Diese Evaluationen sind insofern *nicht herrschaftsfrei*, weil die Überprüfungen und Kontrollen für die betroffenen Institutionen und Personen von existenzieller Bedeutung sind, ihre Messergebnisse zwar protokolliert, die gefundenen Ressourcenmängel aber nicht durch entsprechende Mittelzuweisungen ausgeglichen werden.

Forschung - auch Forschung in der Musikpädagogik - findet also immer in einem gesellschaftlichen Raum statt, ihre Fragestellungen und Ergebnisse werden vor dem Hintergrund gesellschaftlicher Interessen verwertet oder sogar in Abhängigkeit von ihnen formuliert. Wir sollten uns insofern als Forscherinnen und Forscher in der Musikpädagogik immer wieder bewusst sein: Der Erkenntniszuwachs, den unsere Forschung erbringt, muss *für diejenigen, um derentwillen sie betrieben wird, in seinem Nutzen nachvollziehbar und von Bedeutung sein* (Hermann J. Kaiser).

Hinweis: Aus skalierungstechnischen Gründen sind die Textbestandteile in einigen Grafiken der Beiträge von Naacke und Lothwesen nicht optimal lesbar.

Mixed Methods in der Evaluationsforschung

Das Verhältnis zwischen Qualität und Quantität in der Wirkungsanalyse

1 Einleitung

Evaluation, verstanden als Wirkungsmessung, gewinnt eine rasch wachsende Bedeutung in sehr unterschiedlichen Politikfeldern, in den letzten Jahren auch und gerade in der Bildungsforschung: Die großen ländervergleichenden Schulstudien wie TIMMS oder PISA stellen Evaluationen ganzer Bildungssysteme im internationalen Vergleich dar (z.B. Prenzel et al. 2007; Bundesministerium für Bildung und Forschung 2001), bildungspolitische Programme wie die nordrhein-westfälische Initiative >Jedem Kind ein Instrument< werden mit umfangreichen (und teuren) Evaluationsprogrammen begleitet (Kranefeld 2009), in den Hochschulen wird im Zuge der Bologna-Reformen und der immer stärker an einer Kultur der „*accountability*“ orientierten permanenten Studienreform die Bewertung von Lehrveranstaltungen und die evaluationsgestützte Verbesserung der Qualität von Lehre und Studium zu einer Selbstverständlichkeit, die auch gesetzlich abgesichert ist (Landeshochschulgesetze - eine Übersicht findet sich unter http://www.hof.uni-halle.de/steuerung/lhg_uebersicht.htm).

Eine solche Verwissenschaftlichung politischen Handelns weckt Hoffnungen auf eine stärkere Versachlichung und Rationalität solcher Maßnahmen; sie führt natürlich auch zu der Frage, ob und in welchem Umfang die Sozialwissenschaften mit ihrem Methodeninstrumentarium die an sie gerichteten Anforderungen erfüllen können. Evaluationen untersuchen die Wirkungen spezifischer Maßnahmen – das bedeutet, dass es hier immer um die Analyse von Kausalbeziehungen gehen muss. Unter erkenntnistheoretischer und methodologischer Perspektive ist Kausalität ein äußerst komplexer Gegenstand, bei der Aufdeckung und Beschreibung von Kausalbeziehungen gibt es zahlreiche Probleme, Fehler- und Irrtumsmöglichkeiten. Dabei ist es Konsens, dass eine adäquate Wirkungsmessung quantitative Methoden und eine (quasi)experi-

mentelle Logik erfordert, weil nur standardisierte Messungen eine objektive Feststellung zulassen, ob die für den Erfolg einer Maßnahme relevanten Kriterien erreicht wurden, und nur der in (quasi)experimentellen Untersuchungen mögliche kontrollierte Vergleich zwischen Situationen mit und ohne Interventionen kausale Fehlschlüsse zu vermeiden hilft.

Allerdings bringt eine Konzentration auf quantitative Methoden auch Gefahren im Hinblick auf die Validität der Evaluationsdaten mit sich, wie wir im Folgenden anhand einer ausführlichen Diskussion von Stärken und Schwächen qualitativer *und* quantitativer Forschungsmethoden zeigen wollen. So können in ausschließlich quantitativen Erhebungen kaum die Itemverständnisprobleme der Befragten oder auch deren möglicherweise sozial erwünschtes Antwortverhalten deutlich werden und Nebeneffekte von Interventionen, von denen der Forscher keine Kenntnis hat und die damit auch nicht operationalisierbar sind, können oft gar nicht in den Blick geraten. Es lohnt sich deshalb, hier aktuelle Diskussionen in den Sozialwissenschaften über die Bedeutung von „Mixed Methods“ zu berücksichtigen: Durch eine Kombination qualitativer und quantitativer Methoden in einem Forschungsdesign wird es möglich, Schwächen der beiden Methodentraditionen durch die Stärken der jeweils anderen Tradition auszugleichen.

In unserem Beitrag werden wir in einem ersten Schritt kurz auf die aktuelle Debatte um *Mixed Methods* eingehen. Danach werden wir Validitätsbedrohungen und Methodenprobleme qualitativer und quantitativer Forschung bei der Untersuchung von Kausalbeziehungen diskutieren und zeigen, wie die Integration von Methoden dabei helfen kann, solche Methodenprobleme zu entdecken und ggf. zu überwinden. Abschließend werden wir – auch anhand eines eigenen empirischen Beispiels – zeigen, dass die Integration von qualitativen Erhebungsverfahren in eine quantitativ angelegte Untersuchung das Potential bietet, Messprobleme standardisierter Instrumente zu identifizieren und partiell zu überwinden und die „*outcomes*“ von Interventionen und vor allem deren unbeabsichtigte und nicht antizipierte Nebeneffekte zu explizieren.

2 Der Methodenstreit in der empirischen Sozialforschung

Die Forschungsmethodik der Sozialwissenschaften ist seit vielen Jahren durch einen Streit zwischen zwei Traditionen geprägt, die mit jeweils unterschiedlichen Daten und Auswertungsmethoden arbeiten und in denen sich verschiedene Modelle des sozialwissenschaftlichen Forschungshandelns und differierende Qualitätskriterien für gute Forschung entwickelt haben (vgl. Kelle 2008).

Die *quantitative Tradition* orientiert sich am Modell der experimentellen Naturwissenschaften und arbeitet vorzugsweise mit standardisierten Daten, weil nur solche Daten ohne eine weitere Bearbeitung mit statistischen Verfahren analysiert werden können. In dieser Tradition wird üblicherweise ein Modell des sozialwissenschaftlichen Forschungsprozesses verfolgt, bei dem empirische Untersuchungen als Prozesse deduktiver Theoriebildung und anschließenden Theorieprüfung mit Hilfe quantifizierbarer Informationen verstanden werden. In Abgrenzung zu diesem hypothetiko-deduktiven („HD“-)Modell hat sich in den Sozialwissenschaften ein zweiter Traditionsstrang entwickelt, der seine Wurzeln in der geisteswissenschaftlichen Hermeneutik einerseits und der modernen Ethnographie andererseits hat und der den Forschungsprozess als Erkundung eines relativ fremden Terrains versteht. Die in solcher qualitativen Forschung gesammelten Daten sind unstandardisiert und oft auch kaum standardisierbar, das Datenmaterial bilden Texte (und in jüngerer Zeit zunehmend Bilder), die mit interpretativen Verfahren einer sorgsam hermeneutischen Analyse unterzogen werden müssen.

Das Verhältnis zwischen diesen beiden methodologischen Traditionen ist schon lange angespannt und von wechselseitiger Kritik und Abgrenzung gekennzeichnet. Oftmals ist die These vertreten worden, dass qualitative und quantitative Forschungsmethoden notwendigerweise logisch inkompatibel und inkommensurabel seien, weil sie auf jeweils verschiedenen, miteinander unvereinbaren erkenntnistheoretischen Grundpositionen aufbauen würden (vgl. dazu etwa Lamnek 1995, Seite 253; Guba, Lincoln 1988, Seite 93). Allerdings haben seit den 1930er Jahren zahlreiche klassische und neuere sozialwissenschaftliche Untersuchungen, die sowohl zu empirisch neuen und überraschenden Einsichten geführt als auch oft die Theorieentwicklung des Faches außerordentlich stimuliert haben, qualitative und quantitative Forschungsmethoden parallel und gemeinsam in einem Forschungsdesign genutzt. Hierzu zählen etwa die für die Arbeitslosigkeitsforschung paradigmatische „Marienthalstudie“ (Jahoda, Lazarsfeld, Zeisel 1933/1982), die „Hawthorne Study“ (Roethlisberger, Dickson 1939), die die Entwicklung der Industriesoziologie stark beeinflusst hat, die Studie von Festinger, Riecken und Schachter über Weltuntergangskulte (1956), welche die Theorie kognitiver Dissonanz (mit)begründet hat oder Zimbardos bekanntes „Gefängnisexperiment“ über die Folgen institutioneller Deindividuation (Zimbardo 1969).

Aus Bereichen heraus, die ein besonderes Interesse an der praktischen Anwendung sozialwissenschaftlicher Forschungsmethoden haben, wie insbesondere die erziehungswissenschaftliche Forschung, hat sich seit Ende der

1980er Jahre ein methodologischer Diskurs entwickelt, bei dem auf den forschungspraktischen und theoretischen Ertrag einer Kombination von Verfahren beider Traditionen hingewiesen wird – dabei hat sich (vor allem im anglo-amerikanischen Sprachraum) mittlerweile bereits eine „Mixed Methods Bewegung“ etabliert mit eigenen Konferenzen, Handbüchern und Zeitschriften (vgl. etwa Tashakkori und Teddlie 2003; Greene 2007; Kelle 2006, 2008; Creswell 2009; insbesondere auch die nun im dritten Jahr erscheinende Zeitschrift „*Journal of Mixed Methods Research*“). Seit den 1980er und insbesondere in den 1990er Jahren wurde eine wachsende Anzahl von sozialwissenschaftlichen Forschungsprojekten durchgeführt, in denen qualitative und quantitative Methoden miteinander kombiniert wurden - in den deutschen Sozialwissenschaften hat in den 1990er Jahren insbesondere der Sonderforschungsbereich 186 der DFG Pionierarbeiten geleistet (vgl. Kelle 2008, S. 231 ff.; Kluge, Kelle 2001; Heinz, Marshall 2003; internationale Überblicke über Mixed Methods Studien in verschiedenen Disziplinen finden sich etwa bei Tashakkori, Teddlie 2003, S. 491 ff. oder bei Greene 2007). Aus der Mixed Methods Bewegung liegen mittlerweile eine Fülle von Publikationen mit Vorschlägen zur Integration qualitativer und quantitativer Methoden in konkreten Forschungsdesigns vor (ein Überblick etwa bei Creswell et al. 2003; Seipel, Rieker 2003, S. 236 ff.; Greene 2007; Kelle 2008 S. 282 ff.).

Ist der alte Methodenstreit damit obsolet geworden? Ein genauerer Blick auf aktuell laufende Diskussionen zeigt, dass sich die Sache keinesfalls so einfach verhält. Die vorwiegend im anglo-amerikanischen Raum geführte Debatte über Mixed Methods hat noch eine große Anzahl von Problemen zu lösen. Der größte Teil der vorliegenden Arbeiten beschränkt sich auf methodisch-praktische Fragen der Gestaltung von Forschungsdesigns, wobei tiefer liegende Probleme der Relevanz und vor allem Gültigkeit von Forschungsergebnissen oft vernachlässigt werden. Dabei beklagen die Diskutanten oft die vorherrschende begriffliche Unklarheit und eine nach wie vor uneinheitliche Terminologie (Tashakkori, Teddlie 2003). Und Alan Bryman, einer der Gründerväter der Mixed Methods Bewegung, zeigte in einem systematischen Review-artikel über Arbeiten, die mit einem qualitativ-quantitativen Methodenmix arbeiten, dass in solchen Studien qualitative und quantitative Methoden oft nur nebeneinander eingesetzt, aber nicht wirklich miteinander verbunden oder aufeinander bezogen würden (Bryman 2007).

Die Verdrängung der wissenschaftstheoretischen Argumente und Auseinandersetzungen (unter deren Überbetonung die alte Methodendebatte in den Sozialwissenschaften gelitten hatte) zugunsten einer reinen Forschungsprag-

matik führt also ebenso in eine Sackgasse wie der nie gelöste Methodenstreit (vgl. Kelle 2008, Seite 26ff.). Dagegen weist ein von Vertretern der Mixed Methods Bewegung häufig vorgebrachtes Argument, dass eine Methodenkombination sinnvoll ist, um die *Stärken und Schwächen* quantitativer und qualitativer Forschung auszugleichen (vgl. etwa Johnson, Turner 2003, S. 299) in die richtige Richtung. Das bedeutet aber auch: Es müssen zuerst die Schwächen qualitativer und quantitativer Forschung genau beschrieben und geklärt werden, damit dann in einem zweiten Schritt diese Schwächen durch die Stärken der jeweils anderen Tradition ausgeglichen werden können – erst dann kann eine sinnvolle Methodenkombination stattfinden. Arbeiten, die so vorgehen, sind nun aber auch in der gegenwärtigen Debatte über Mixed Methods nicht immer leicht zu finden, so dass man in der entsprechenden Literatur kaum methodologische Regeln findet, die eine Entscheidung darüber zulassen, an welchen Punkten des Forschungsprozesses für welche (Teil)fragen welche Methoden einzusetzen sind – so dass bei der Anwendung eines Mixed Methods Designs bei der Auswahl von Methoden letztlich *ad hoc* Entscheidungen getroffen werden müssten. Die Frage nach den Schwächen und Grenzen beider Traditionen, die im aktuellen Mixed Methods Diskurs oft ausgeklammert bleibt, wurde im klassischen Methodenstreit der Sozialwissenschaften aber zumindest formuliert und diskutiert. Das Problem dieser Debatten war nur, dass beide Lager sich stets auf Schwächen der Gegenseite konzentriert haben, während man die von dort aus vorgebrachte Kritik an der eigenen Tradition ignoriert oder ihr auch manchmal durch sprachliche Kunstgriffe auszuweichen trachtet (indem man etwa versucht, der Besorgnis über das Problem der mangelnden Generalisierbarkeit von Fallstudien durch sprachliche Manipulation am Begriff wissenschaftlicher Gültigkeit zu begegnen, etwa den Begriff der „Generalisierbarkeit“ durch den der „Übertragbarkeit“ von Ergebnissen zu ersetzen, vgl. Lincoln, Guba 2000 und zur Kritik an diesen Positionen Kelle 2008, S. 37 ff.).

Die Diskussion um Mixed Methods eröffnet aber die Möglichkeit, die Methodendebatte aus dieser Sackgasse zu führen, indem Kritik als Ressource genutzt wird, um die Schwächen bestimmter Ansätze zu erkennen, zu bearbeiten und (ggf. durch die Verwendung von Methoden aus einer alternativen Tradition) zu beheben. Auf diese Weise kann sich das methodische Instrumentarium der Sozialwissenschaften durch kritische Kooperation weiter entwickeln, was Streit nicht ausschließt, aber eine selbstkritische Einstellung erfordert, die bei einer ernsthaften Bestandsaufnahme der methodischen Probleme und Grenzen des eigenen Vorgehens und bei der Verbesserung methodischer Werkzeuge hilft. Auf dieser Grundlage kann das Querschnittsfach „Methodologie“ seine

klassische Aufgabe erfüllen, „Fehlertheorien“ zu formulieren und zu klären, welche Schwächen konkrete Methoden bei der Untersuchung spezifischer sozialwissenschaftlicher Gegenstände jeweils haben und wo dies zu blinden Flecken bei der empirischen Forschung führt.

Im Folgenden möchten wir diese Strategie auf das Feld der Evaluationsforschung beziehen, indem wir herausarbeiten,

1. wo die spezifischen Schwächen und Probleme qualitativer und quantitativer Methoden bei der empirischen Analyse kausaler Beziehungen liegen, und
2. wie diese Schwächen durch die Einbeziehung von Methoden der jeweils anderen Tradition bearbeitet und vielleicht auch ausgeglichen werden können.

3 Kausale Analyse und die Stärken und Schwächen qualitativer und quantitativer Forschung

Evaluationsforschung, insbesondere verstanden als „strenge Wirkungsevaluation“ („*rigorous impact evaluation*“, vgl. bspw. White 2006) untersucht Kausalzusammenhänge: eine bestimmte Intervention oder Maßnahme soll „wirken“, das heißt sie soll einen beschreibbaren und messbaren „Effekt“ haben. Dabei können die Vorstellungen, die diejenigen von diesen Wirkungen haben, die eine Intervention planen und durchsetzen, mehr oder weniger genau und spezifisch sein, so dass Evaluatoren oft erst gemeinsam mit den Gestaltern von *policies* genau heraus arbeiten müssen, wie man Indikatoren für Wirksamkeit in dem konkreten Untersuchungsbereich entwickelt, das heißt, woran man genau erkennen kann, dass eine bestimmte Maßnahme wirkt. Bei einer strengen Wirkungsmessung setzen wissenschaftliche Evaluatoren und politische Entscheider oft gleichermaßen ihr Vertrauen auf standardisierte quantitative Messverfahren und das nicht ohne Grund: in der quantitativen Methodentradiation gibt es seit den 1950er Jahren eine lange Tradition in der Diskussion kausaler Zusammenhänge und in der Bearbeitung der dort auftretenden Probleme (vgl. bspw. Lazarsfeld 1955; Simon 1954; Blalock 1985), während dieses Themenfeld in der qualitativen Forschung bisher eher gemieden wurde. Dort wird der Kausalitätsbegriff bisher lediglich in den Arbeiten von Anselm Strauss (Strauss/Corbin 1990, 101) und ansatzweise in Konzepten, die auf das Verhältnis von „Erklären“ und „Verstehen“ eingehen (vgl. bspw. Lamnek 1995, 74), verwendet. Weil darüber hinaus die Standards quantitativer Erhe-

bungen mit ihren Gütekriterien der Objektivität und Reliabilität sowie der Repräsentativität der Stichprobenziehung auch für viele Auftraggeber von Evaluationsprojekten starke und nachvollziehbare Argumente bieten, ist eine Dominanz quantitativer Verfahren und eine eher marginale Bedeutung qualitativer Erhebungsmethoden im Rahmen der Evaluationsforschung durchaus erklärbar. Dabei sind qualitative Methoden unabdingbar, um die grundlegenden Defizite, die quantitative Erhebungen bei der Untersuchung von Kausalbeziehungen aufweisen, aufzudecken und zumindest partiell zu überwinden.

3.1 *Die Stärken quantitativer und experimenteller Ansätze*

Aber in welchen Punkten sind quantitativ-standardisierte und quasi-experimentelle Verfahren den qualitativen Methoden tatsächlich überlegen? Folgt man der Argumentation eines wissenschaftsphilosophischen Fallibilismus, wie er in den Arbeiten von Charles Sanders Peirce und Karl Popper (für die Sozialwissenschaften von Hans Albert) begründet und dann in verschiedenen (auch liberaleren) Fassungen von Imre Lakatos, Nicolas Rescher und Larry Laudan weiterentwickelt wurde (z.B. Chalmers 2007), unterscheidet sich wissenschaftliche Forschung von Alltagsstrategien des Wissenserwerbs vor allem durch ihre besondere methodische Kontrolle, mit der versucht wird, typische Fehler zu vermeiden, die der *common sense* bei einer Formulierung und Überprüfung von allgemeingültigen Aussagen, bei der Untersuchung empirischer Gegenstände und bei der Formulierung von Schlussfolgerungen leicht begeht. Zur konkreten Sicherung methodischer Kontrolle in der Forschungspraxis kann Methodologie (als Querschnittsdisziplin zwischen Sozial- und Naturwissenschaften) einerseits „Fehlertheorien“ formulieren, also Theorien darüber, welche Fehler bei der Untersuchung der empirischen Realität auftreten können, und andererseits Vorschläge ausarbeiten, wie diese Fehler vermieden werden können.

Fehlermöglichkeiten in der Evaluationsforschung können bspw. darin bestehen, dass entweder eine Veränderung konstatiert wird, wo tatsächlich gar keine Veränderung stattgefunden hat, oder aber dort wo eine (möglicherweise nicht-intendierte) Veränderung stattgefunden hat, diese Veränderung übersehen wird.

Der erste Fehler wird leicht dort geschehen, wo keine angemessenen (oder falsche) Beobachtungsverfahren und Messinstrumente eingesetzt worden sind. Wahrnehmungs- und Kognitionspsychologen haben in ihren empirischen Forschungen gezeigt, wie stark die menschliche Wahrnehmung dazu tendiert, Einzelbeobachtungen, die mit Vorannahmen oder auch Stereotypen konsistent

sind, zu registrieren und Gegenevidenz (selbst wenn sie häufig und massiv auftritt), zu übersehen (z.B. Kahneman 1982/2008; Festinger 1957). Diese kognitive Tendenz, bestätigende Evidenz für die eigenen Annahmen zu stark und Falsifikatoren und Anomalien zu gering zu bewerten, kann zum „Rosenpicken“ führen, bei dem vermeintliche Belege für die Wirksamkeit der eigenen Maßnahme zusammengesucht und präsentiert werden.

Beliebt ist dieses Vorgehen etwa bei der Präsentation von pseudo- und alternativmedizinischen „Wunderkuren“ oder im Marketing, wo gerne unterstützende und bestätigende Berichte einiger „begeisterter“ Patienten oder Kunden präsentiert werden. Im Kontext der „evidenzbasierten Medizin“ steht deshalb die einfache Präsentation von Fallgeschichten auf der niedrigsten Stufe wissenschaftlicher Evidenz. Aber auch im Bereich von sozialpolitischen Interventionen ist ein solches Vorgehen – Beleg der Wirksamkeit durch Einzelfallberichte – häufig. Was hier dann übersehen wird: möglicherweise stellt sich der beobachtete Effekt nur unter ganz bestimmten (sehr eingeschränkten) Bedingungen ein, die gar nicht genau bekannt sind, während die Intervention für große Teile der betreffenden Zielgruppe völlig wirkungslos ist.

Die Wahrnehmung (oder die Leugnung) von Veränderungen kann auch die Folge einfachen Wunschdenkens (vgl. dazu etwa Babad, Katz 1991 zur Bedeutung von „*wishful thinking*“) sein.

Diese Art der Wahrnehmungsverzerrung hat einer der Begründer der Logikerschule von Port Royal, Antoine Arnauld bereits 1662 mit seinen Aphorismen aufgespießt: „Wir beurteilen die Dinge, nicht das zugrunde legend, was sie an und für sich sind, sondern das, was sie in der Beziehung zu uns sind.“ (Arnauld 1972/1662, S. 253) Und: „Wenn sie jemanden lieben, ist er von jeder Art Fehler befreit. Alles, was sie wünschen, ist gerecht und leicht, alles, was sie nicht wünschen, ist ungerecht und unmöglich, ohne dass sie irgendeinen Grund für alle diese Urteile vorzubringen vermöchten, außer der Leidenschaft, von der sie besessen sind.“ (ebd., S. 254)

Neben der Unter- oder Überschätzung von (erwünschten oder unerwünschten) Veränderungen von Interventionen gibt es Fehlermöglichkeiten, die den Kausalnexus zwischen Intervention und Veränderung betreffen: wenn die erwünschte Veränderung nicht wegen der Intervention, sondern aufgrund anderer Ursachen eintritt, kann eine kausale Verbindung zwischen Intervention und Veränderung unterstellt werden, die gar nicht existiert. Eine solche kausale Fehlattribution wird (nicht nur) in der methodologischen Literatur häufig beschrieben, es handelt sich um das klassische Fehlurteil des „post hoc, ergo propter hoc“, das in nahezu allen Handlungsfeldern, in denen Interventionen stattfinden, eine Rolle spielen kann und das unter verschiedenen begrifflichen Etiketten diskutiert wird:

In der Medizin gehört das Wissen um „Spontanremissionen“ zur Folklore: zahlreiche Krankheiten heilen auch ohne medizinische Interventionen („Ein Schnupfen dauert mit

Behandlung eine Woche, ohne Behandlung sieben Tage“). Das bedeutet aber auch: bei jeder medizinischen Intervention (auch in größeren Gruppen), der eine Heilung oder Veränderung folgt, muss die Frage gestellt werden, ob die Heilung nicht auch ohnehin geschehen wäre. Für den Bereich der Sozial- und Erziehungswissenschaften sprechen Campbell und Stanley in ihrer klassischen Arbeit über „Experimentelle und quasi-experimentelle Designs“ von „Reifungseffekten“ – eine bestimmte Lehrmethode wird vielleicht vor allem deswegen als erfolgreich angesehen, weil Schüler aufgrund normaler kognitiver Reifungsprozesse einen Kompetenzzuwachs erfahren (vgl. Campbell, Stanley 1963).

In der Tradition der quantitativ-standardisierten Forschung wurde eine Reihe von Vorsichtsmaßnahmen vorgeschlagen und ausgearbeitet, um diese einfachen Beobachtungsfehler zu beherrschen und zu begrenzen:

- So kann man Fehler, die durch „*wishful thinking*“ entstehen, das heißt Wahrnehmungsverzerrungen, die dazu führen, dass man dort eine positive Veränderung sieht, wo eine solche Veränderung gar nicht stattgefunden hat, oder die eine tatsächlich stattgefundene Veränderung übertreiben und überzeichnen, durch Verfahren einer objektiven (das heißt: beobachterunabhängigen) standardisierten Beobachtung und Messung zu begrenzen versuchen. Ein zuverlässiger empirischer Befund über eine bestimmte Veränderung liegt demnach eben nur dann vor, wenn mehrere neutrale Beobachter (auch solche, die die Vorannahmen und Wünsche des Untersuchers nicht teilen) bei dem Vorliegen derselben Daten und Befunde übereinstimmend zu demselben Urteil über diese Veränderung gelangen. In den Sozialwissenschaften ist die Methode der standardisierten Befragung ein Mittel, um Wahrnehmungsverzerrungen und Wunschdenken von Untersuchern zu vermeiden: Täuschungen darüber, ob Befragte an bestimmten Stellen ein Kreuz gemacht haben, sind selten. Demgegenüber lassen die interpretativen und hermeneutischen Verfahren, die in der qualitativen Sozialforschung angewendet werden, *prima facie* stets Interpretationsspielräume bei der Auswertung des Datenmaterials zu – die hierdurch aufgeworfenen Probleme müssen dann durch sehr komplexe und aufwändige Methoden (bspw. in Gruppenauswertungen oder durch eine EDV-gestützte synoptische Analyse des Materials) bearbeitet werden.
- Das Risiko des „Rosinenpickens“ lässt sich dadurch vermeiden, dass die Anzahl der empirischen Beobachtungen systematisch erhöht wird, indem also etwa alle Betroffenen einer Maßnahme befragt werden, oder indem aus dieser Gruppe eine (Zufalls)Stichprobe *lege artis* gezogen wird. Wenn alle diese Fälle dann auch untersucht werden, können Gegenbeispiele und *negative cases* (d.h. Fälle, bei denen keine Veränderung stattgefunden hat), nicht mehr ohne weiteres übersehen und ausgeblendet werden. Die Stichproben-

größe, die notwendig ist, um Gegenbeispiele zu erfassen, hängt nun auch von der Heterogenität des Untersuchungsfeldes (oder, technisch gesprochen, von der Varianz des untersuchten Merkmals) ab - da qualitative Erhebungs- und Analyseverfahren sehr aufwändig sind, lassen sich hiermit allerdings immer nur kleine Fallzahlen bearbeiten, so dass in dieser Tradition das Risiko einer einseitigen und verzerrenden Fallauswahl im Allgemeinen höher ist als in der quantitativen Forschung.

- Schließlich bieten experimentelle Interventionen ein sehr brauchbares Werkzeug, um klassische kausale Fehlschlüsse zu erkennen und zu vermeiden: Kausale Fehlschlüsse, bei denen Spontanremissionen oder Reifungsprozesse übersehen wurden und deshalb eine Veränderung, die ohnehin von selber eingetreten wäre, auf eine Intervention zurückgeführt wird, obwohl diese Intervention hierfür gar nicht ursächlich war, lassen sich durch die Bildung von Kontrollgruppen (also von Gruppen, bei denen keine Intervention stattfindet), vermeiden. Auch andere Hypothesen über alternative Ursachen für die interessierenden Veränderungen lassen sich durch experimentelle Kontrollen prüfen. Nun sind bei der Durchführung politischer Interventionen und sozialtechnologischer Maßnahmen nur selten strenge experimentelle Designs einsetzbar (eine zufällige Aufteilung von Betroffenen auf Versuchs- und Kontrollgruppen ist etwa oft aus politischen, praktischen oder ethischen Gründen gar nicht möglich), jedoch gibt es in der sozialwissenschaftlichen Methodenforschung seit den bahnbrechenden Arbeiten von Campbell und Stanley (1963) eine recht umfangreiche Literatur über verschiedene Möglichkeiten der „quasi-experimentellen“ Kontrolle.

Die quantitative Methodentradiation stellt eine Reihe von ausformulierten Fehlertheorien (etwa inferenzstatistische Theorien des Stichprobenfehlers) und darauf aufbauende Verfahren zur Fehlererkennung und Fehlervermeidung zur Verfügung, deren routinemäßige Anwendung in der Praxis der empirischen Sozialforschung (zumindest dort, wo *lege artis* gearbeitet wird) verbreitet ist. Qualitative Verfahren werfen gerade in Bereichen, in denen die quantitative Tradition Lösungen anbietet, oft besondere Methodenprobleme auf: so ist hier häufig allein aus forschungspraktischen Gründen die Prüfung, ob verschiedene Forscher unabhängig voneinander zu denselben Ergebnissen kommen würden (Kelle 2008, 28ff.), nur schwer möglich. Fallzahlen in der qualitativen Forschung (aufgrund des besonderen Aufwands der hier eingesetzten Erhebungs- und Auswertungsmethoden) sind oft sehr klein, so dass sich sofort Fra-

gen nach der Verallgemeinerbarkeit der hier gewonnenen Befunde aufdrängen (Kelle 2008, S. 231).

3.2 *Methodenprobleme und Validitätsbedrohungen quantitativer Untersuchungen*

Angesichts dieser Situation stellt sich natürlich die Frage, warum es dann überhaupt sinnvoll sein kann, qualitativ zu forschen. Nun haben auch quantitative Methoden ganz spezifische Beschränkungen und Methodenprobleme, unter anderem dort, wo Kausalität als sozialwissenschaftliches Konzept verwendet wird, das heißt dort, wo die untersuchten Ursache- und Wirkungsphänomene *soziale Handlungen* sind. Interessanterweise stammt der Begriff *causa* historisch auch gar nicht aus einem naturwissenschaftlichen Kontext, sondern wurde zuerst in sozialen Zusammenhängen in enger Verbindung mit dem Konzept des Handelns verwendet, wie R.G. Collingwood in einer sprachgeschichtlichen Untersuchung gezeigt hat (vgl. Collingwood 1937/1938). Begriffe, die auf Ursachen verweisen, wie „bewirken“, „verursachen“, „beeinflussen“ oder „führen zu“ werden seit der Antike (und bis heute nicht nur im alltäglichen Sprachgebrauch, sondern auch in vielen wissenschaftlichen Kontexten, die sich auf menschliches Handeln beziehen, etwa in den Rechtswissenschaften) nicht primär auf Naturprozesse, sondern auf freie und intendierte Handlungen von verantwortlichen Akteuren bezogen. Soziales Handeln kann somit in der ursprünglichen Bedeutung des Begriffs *causa* durchaus als (wenn auch nicht deterministisch) beeinflusst oder verursacht verstanden werden. Folgt man einer bereits bei Aristoteles getroffenen Unterscheidung, besteht die *causa* einer Handlung aus der *causa ut*, einer bestimmten Absicht eines Akteurs, und der *causa quod*, einer bestimmten Situation oder einem bestimmten Sachverhalt. Diese Differenzierung lässt sich auch in modernen handlungstheoretischen Konzeptionen finden: Alfred Schütz zufolge erfordert das Verstehen und die Erklärung sozialen Handelns, dass einerseits „um-zu Motive“, d.s. Ziele, Motive oder Absichten des Handelnden (die *causa ut*) und andererseits „weil-Motive“, d.s. Situationsmerkmale (die *causa quod*) einbezogen werden (vgl. Schütz 1971). Natürlich ist aber ein solches Verstehen sozialen Handelns nicht möglich ohne eine gewisse Kenntnis allgemeiner Regeln, mit denen sich Handlungsziele (die „um-zu-Motive“) und Handlungsbedingungen (die „weil-Motive“) miteinander verknüpfen lassen. Aus der Perspektive nahezu aller aktuellen Handlungstheorien muss dabei davon ausgegangen werden, dass die *Handlungsziele* von Akteuren plastisch und flexibel sein können und sich an veränderte *Handlungsbedingungen* anpassen, dass Handlungsbedingungen

wiederum oft variabel sind und von kreativen, entscheidungsfähigen Akteuren veränderbar und dass schließlich jene *Handlungsregeln*, mit denen Handlungsbedingungen und Handlungsziele miteinander verbunden werden, nicht universell sind, sondern kulturspezifisch und lokal gelten und soziokulturellem Wandel unterliegen.

Das klassische Kausalmodell der quantitativ orientierten Evaluationsforschung kann mit solch einer Flexibilität und Kontingenz von Handlungszielen, Handlungsbedingungen und Handlungsregeln oft nur schwer umgehen. Das Standardmodell nimmt hier die folgende Form an: Zur Überprüfung der Hypothese, dass eine Intervention einen bestimmten *outcome* erzeugt, müssen die Intervention als unabhängige Variable und der *outcome* als abhängige Variable operationalisiert werden. Eine kausale Wirkung wird dann angenommen, wenn die Veränderung der unabhängigen Variablen (die idealerweise durch ein experimentelles Design erfolgt, in dem eine randomisierte Zuordnung der Untersuchungseinheiten zu den Ausprägungen der unabhängigen Variablen erfolgt) auch eine messbare Änderung der abhängigen Variablen nach sich zieht (Grafik 1).

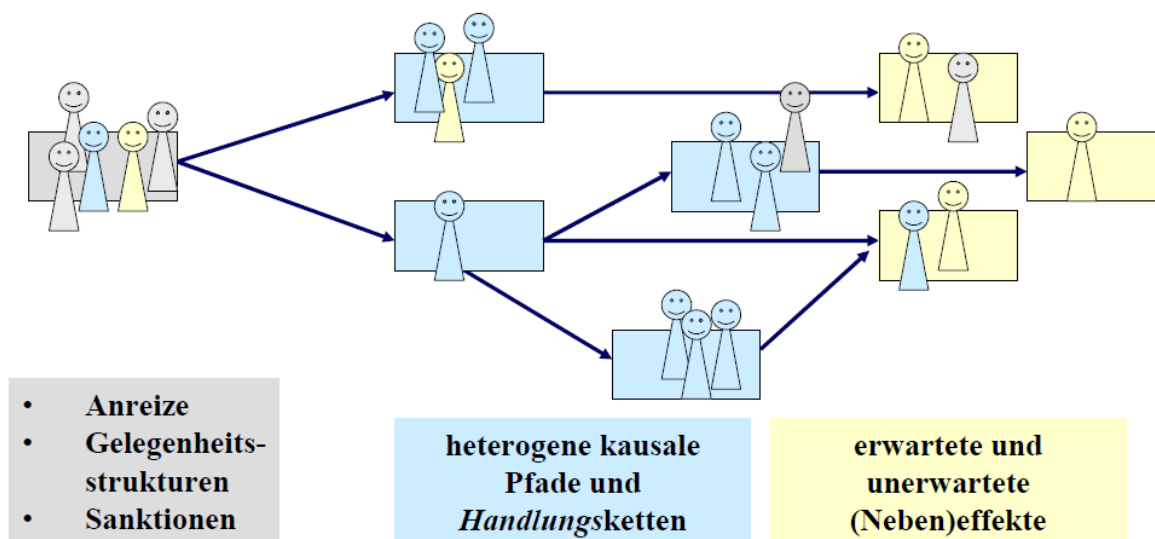


Grafik 1: Intervention und Outcome

Bildungs- und sozialpolitische Interventionen¹ involvieren nun komplexes soziales Handeln und erfordern lange instrumentelle Ketten, die zahlreiche Einzelhandlungen einer großen Anzahl individueller Akteure miteinander vernetzen, welche jeweils verschiedene Handlungsziele und einen unterschiedlichen kulturellen Hintergrund haben. Die verschiedenen Einzelhandlungen generieren dabei große Mengen verschiedener Handlungsergebnisse, die erst zusammengenommen, im Aggregat, darstellbar etwa durch statistische Mittelwertbildung, als Gesamt-*outcome* betrachtet werden können (siehe Grafik 2). Das grundlegende Problem besteht nun darin, dass der größte Teil des Netzwerks von Handlungsketten, die sich durch die Interventionsmaßnahme her-

1 Etwa: die Einführung des Programms „Jedem Kind ein Instrument“ an nordrhein-westfälischen Grundschulen oder auch: die Einführung eines einheitlichen Systems gestufter Studiengänge und Abschlüsse in der Universitätsausbildung in verschiedenen Staaten, die sich darauf verständigen.

ausbilden, niemals völlig durch jene Akteure kontrolliert werden kann, die die Intervention planen und anordnen. Durch eine Interventionsmaßnahme lassen sich nur bestimmte *Handlungsbedingungen* der Akteure vor Ort *direkt* beeinflussen, nicht deren Handlungen selber. Eine Intervention kann Anreize schaffen (etwa: *ökonomische Anreize für Musikschullehrer*), sie kann Gelegenheitsstrukturen institutionalisieren (etwa: *Angebote an Musikunterricht in Grundschulen* usw.) oder sie kann administrative Sanktionsdrohungen (etwa: die Kürzung staatlicher *Mittel*) verhängen.



Grafik 2: Komplexe sozialpolitische Intervention mit zahlreichen Akteuren

Hierdurch werden Akteure aber nicht mechanisch gezwungen, sondern nur motiviert, im Sinne des gewünschten *outcome* zu handeln, weil dies bestimmten Handlungszielen dient (bspw.: *in den Genuss ökonomischer Vorteile zu gelangen, Qualifikationszertifikate zu erwerben, oder die Kürzung finanzieller Mittel vermeiden*), die ihnen von denjenigen, die die Intervention planen, mehr oder weniger sinnvoll unterstellt werden können. Es ist allerdings durchaus möglich, dass Akteure, die die Interventionen umsetzen sollen, andere und konkurrierende Handlungsziele verfolgen. Akteure sind in der Regel in lokale kulturelle und institutionelle *settings* eingebunden, in denen bestimmte Handlungsziele und die zu ihrer Erreichung notwendigen Mittel und Wege bekannt sind. In vielen Fällen kennen aber weder diejenigen, die die Intervention planen, noch diejenigen, die sie evaluieren, solche lokalen Organisations- oder Professionskulturen. Zudem können findige Akteure Strategien entwickeln, um die Handlungsbedingungen, die durch die Intervention vorgegeben wurden, in einem Sinn zu beeinflussen, der den Intentionen der sozialpolitisch

Intervenierenden zuwiderläuft. Administrative Sanktionen können umgangen werden, Anreize können konsumiert werden, ohne dass entsprechende Gegenleistungen erbracht werden, Angebote können ignoriert werden, dadurch können Mitnahmeeffekte entstehen u.v.a.m.. Auf diese Weise können Interventionen zahlreiche unbeabsichtigte Effekte nach sich ziehen, einerseits dadurch, dass die an der Intervention Beteiligten und von ihr Betroffenen in lokale kulturelle und institutionelle Kontexte eingebunden sind, die sie mit Handlungszielen und Handlungswissen ausstatten, die Planern und Evaluatoren nicht bekannt sind, und andererseits dadurch, dass die beteiligten Personen entscheidungs- und handlungsfähige Subjekte sind.

3.2.1 Das Problem der übersehenen (Neben)wirkungen

Hiermit gelangen wir zum ersten grundlegenden Problem quantitativer „Monomethodendesigns“ in der Evaluationsforschung: das *Problem der übersehenen (Neben)wirkungen* von Evaluationen. Quantitative Forschung folgt dem Grundprinzip des „hypothetiko-deduktiven Modells“ des Forschungsprozesses. Das bedeutet, dass am Anfang empirischer Forschung stets die Formulierung von präzisen Hypothesen (in der Evaluationsforschung sind dies spezifische, genau ausformulierte Annahmen über konkrete Wirkungen der Interventionen) stehen muss, mit deren Hilfe dann die gemessenen Variablen definiert und die Messinstrumente konstruiert werden. Ohne ein fertiges Messinstrument können standardisierte Daten nicht erhoben werden, was bedeutet, dass die Forscher schon vor der empirischen Untersuchung wissen müssen, welche konkreten Folgen der Intervention sie untersuchen wollen.

Wenn wir allerdings davon ausgehen können, dass die Intervention durch das Handeln entscheidungsfähiger und kreativer Akteure auf der Grundlage von (den Forschern oft gar nicht bekannten) lokalen Wissensbeständen umgesetzt werden (und auch bei dieser Umsetzung in einer Weise verändert werden, die den Intentionen derjenigen, die die Intervention geplant haben, gar nicht entspricht), müssen wir stets mit einer Reihe von (vor einer Intervention nicht antizipierten und vielleicht auch gar nicht antizipierbaren) Nebenfolgen rechnen.

Bemühungen zu einer verstärkten Integration und Inklusion behinderter Schüler in die Regelschule können etwa neue psychosoziale Probleme bei besonders vulnerablen Individuen oder Gruppen aufwerfen: so ist bspw. für stotternde Kinder in Regelschulen das Risiko besonders hoch, während der Schulzeit Opfer von Mobbing, verbalem und physischen Missbrauch durch nichtbehinderte Mitschüler zu werden (vgl. Benecken, Spindler 2004).

Solche unerwarteten Effekte und Folgen können bei einem rein hypothesentestenden Design jedoch überhaupt nicht in den Blick geraten, zu ihrer Erforschung braucht es explorative, qualitative Methoden.

3.2.2 Das Problem der unbekannten Störvariablen

Eng mit dem ersten Problem verbunden ist das *Problem der unbekannten Störvariablen*, die die Umsetzung einer Intervention und die Erreichung der geplanten Wirkungen be- oder verhindern können. Da soziale Interventionen immer durch das Handeln ganz unterschiedlicher Akteure mit verschiedenen Interessen, Zielen und Wissensbeständen umgesetzt werden müssen, erklären sich die Wirkungen von Störvariablen auch oft durch die (mehr oder weniger gezielten) Handlungen der bei der Umsetzung der Intervention beteiligten Akteure und Akteursgruppen.

Ein berühmtes Beispiel aus der Geschichte der Bildungsforschung, das bis heute Gegenstand wissenschaftlicher, politischer und juristischer Auseinandersetzungen ist, zeigt, wie die Intentionen von Forschern und politischen Akteuren durch Betroffene konterkariert werden können (vgl. Coleman 1976, 1977; Hage, Meeker 1988, S27 ff.). Angeregt durch empirische Evidenz, wonach die schlechte Performanz schwarzer Schüler eine Folge ihrer Beschulung in sozial und ethnisch homogenen Schulen war, versuchten US-amerikanische Schulreformer in den 1960er Jahren eine ethnische Desegregation durchzusetzen. Die Erfolge der drastischen Maßnahmen (etwa „busing“, bei dem Schulkinder manchmal länger als eine Stunde von der elterlichen Wohnung zu ihrer neuen Schule transportiert wurden) blieben äußerst beschränkt. Weil viele weiße Mittelschichteltern mit Schulflucht reagierten, erreichten die sozialpolitischen Maßnahmen in manchen Städten sogar das Gegenteil des angestrebten *outcomes*, indem ethnische und soziale Ungleichheiten verstärkt und nicht reduziert wurden.

In der erkenntnistheoretischen Diskussion um das Kausalitätskonzept sind solche Phänomene unter der Bezeichnung „Simpsons Paradoxon“ (Simpson 1951) bekannt. Wir haben es hier – technisch gesprochen – mit einer intervenierenden Variable (der Gegenwehr bestimmter Akteure gegen die Intervention) zu tun, deren Wirkung umso stärker wird, je stärker auch die unabhängige Variable (also die Intervention) wirkt. Die Intervention kann also sehr effektiv sein und gerade deshalb keine messbaren Wirkungen auf der Ebene empirischer Beobachtungen erzielen, weil sie soziale Gegenkräfte auf den Plan ruft. Auch solche Phänomene lassen sich nur mit Hilfe intensiver und aufwändiger Explorationen im Feld, nicht jedoch mit vorab entwickelten Standardmessinstrumenten, in den Blick nehmen.

3.2.3 Das Problem der unbeobachteten Heterogenität

Interventionen sind stets in bestimmten Situationen und mit einem bestimmten Klientel besonders effektiv, und erzielen unter anderen Bedingungen und mit anderen Zielgruppen kaum nennenswerte Erfolge. Oft sind die ansprechbaren und die nicht ansprechbaren Zielgruppen gemischt und können (bereits aus praktischen Gründen) nicht sinnvoll voneinander getrennt werden. Manche dieser intervenierenden Variablen, die die „Ansprechbarkeit“ des Klientels oder die konkreten Rahmenbedingungen der Intervention vor Ort betreffen, sind den Planern der Intervention und den Evaluatoren bekannt, andere nicht.

Ein Programm wie „Jedem Kind ein Musikinstrument“ etwa wird in bestimmten Schulen, bei bestimmten Lehrern und bei bestimmten Schülern besonders starke, in anderen Fällen gar keine oder nur sehr schwache Effekte erzielen. Manche der intervenierenden Variablen lassen sich dabei leichter antizipieren und erfassen als andere: so liegt die Hypothese nahe, dass sich besonders dort starke Effekte erzielen lassen, wo die Eltern das Programm unterstützen und ihre Kinder zu Hause zum Üben bewegen. Diese Unterstützung der Eltern lässt sich mit vertretbarem Aufwand durch standardisierte Items in einem Fragebogen erfassen. Andere Variablen, wie das Engagement von Lehrern, die Unterstützung durch die Schulleitung u.ä. lassen sich weitaus schwerer, teilweise gar nicht, mit quantitativen Methoden untersuchen.

Quantitative Forscher verfügen zwar mit ausgearbeiteten Stichprobentheorien und den Methoden der Stichprobenziehung über einige Werkzeuge, um mit unbeobachtbarer Heterogenität umzugehen: solange es sinnvoll ist, von der klassischen statistischen Normalverteilungstheorie der Fehler auszugehen (das heißt von der Annahme, dass sich verstärkende und hemmende Effekte intervenierender Variablen *grosso modo* ausgleichen), und anzunehmen, dass quantitative Untersuchungen bei genügendem Stichprobenumfang Ergebnisse liefern, die brauchbare Schätzer für Durchschnittseffekte der Intervention abgeben. Nur ist leider ein solches Denken in Mittelwerten nicht immer ohne Probleme: denn die Schwankungsbreite der Effekte (technisch: die Fehlervarianz) kann so hoch werden, dass die Effekte bei nicht sehr umfangreichen Stichprobengrößen nicht mehr signifikant sind (und damit die Intervention als nicht oder kaum erfolgreich gilt, obwohl sie bei bestimmten Gruppen sehr starke Effekte zu erzielen vermag). Und Planer von Interventionen und Evaluatoren sind oft nicht daran interessiert, Minimaleffekte im Durchschnitt zu konstatieren, sondern eher daran, herauszufinden, *wo* (d.h. unter welchen Bedingungen und bei welchem Klientel) eine Maßnahme besonders gut wirkt.

3.2.4 Das Problem systematischer Messfehler und von Methodenartefakten

Das Messparadigma quantitativ-standardisierter Sozialforschung zielt auf eine objektive (d.h. beobachterunabhängige), wiederholbare und möglichst fehlerfreie Erfassung individueller Merkmale (also z.Bsp. der Leistungszuwächse von Schulkindern, oder der Zufriedenheit von Leistungsempfängern), die als Indikatoren für den Erfolg bestimmter Interventionen gelten können. Doch erfassen die hierzu verwendeten Instrumente, etwa Fragebögen, tatsächlich das, was sie messen sollen (Leistung, Zufriedenheit...)? Diese notwendige Frage nach der Validität von Messinstrumenten kann mit den verfügbaren teststatistischen Kontrollprozeduren, welche helfen, die Qualität von Fragebogeninstrumenten zu bewerten, in der Regel kaum beantwortet werden. Diese Prozeduren richten sich nämlich in der Regel stets nur auf die Gütekriterien der *Objektivität* und der *Reliabilität*, die mithilfe statistischer Koeffizienten (etwa anhand von Messwiederholungen) statistisch abgesichert werden können. Demgegenüber ist die „Inhaltsvalidität“ eines Fragebogens (also das Ausmaß, in dem bspw. ein Instrument zur Lehrevaluation wirklich das Kriterium „Lehrqualität“ misst), nicht wirklich rechnerisch erfassbar. Amerikanische Autoren versuchen in jüngerer Zeit diesen Aspekt mit dem Begriff der „*substantive validity*“ (Onwuegbuzie et al. 2007, 117) zu erfassen, welcher primär die Qualität der verwendeten Items und Verständnisdivergenzen betrifft. Um mit Hilfe von Fragebögen valide Ergebnisse zu erzielen, müssen Forscher und Befragte nämlich bspw. kulturelle Bedeutungen miteinander teilen (Cicourel 1974, 29), indem sie Begriffe (wie sie bspw. in einem Fragebogen verwendet werden) übereinstimmend verstehen und interpretieren. Weiterhin müssen beide Gruppen auf dieselben Sprachsysteme und Wissensbestände zurückgreifen können. Lassen unpräzise Formulierungen die Möglichkeit für ein unterschiedliches Begriffsverständnis zu und eröffnen sie Interpretationsspielräume, sind die Ergebnisse nicht mehr eindeutig interpretierbar (Prüfer/ Rexroth 2005, 6). Nur dann, wenn alle Befragten alle Items auf dieselbe Weise und in der vom Forscher intendierten Bedeutung verstehen und alle Begriffe eindeutig interpretierbar sind, können valide Ergebnisse erwartet werden. Weiterhin müssen die Befragten über das grundlegende Wissen verfügen, um die Items adäquat beantworten zu können und hinreichend motiviert sein, den Fragebogen (auch ehrlich) zu beantworten (Fowler 1995, 4).

Außerdem können sich Befragte bei der Beantwortung der Items aus verschiedenen Gründen dafür entscheiden, sozial erwünscht zu antworten oder ihre tatsächliche Einstellung nicht ehrlich preiszugeben (Prüfer/ Rexroth 2005,

4), so dass auch in diesen Fällen die Validität der Ergebnisse beschränkt bleibt.

Aus unserer eigenen Forschung können wir hierzu einige empirische Beispiele berichten. In Untersuchungen über Statuspassagen zwischen Hochschulausbildung und Erwerbstätigkeit in der ehemaligen DDR, bei denen ostdeutsche Hochschulabsolventen kurz nach der Wende befragt wurden, wie sie zu ihrer ersten Arbeitsstelle gelangt sind, neigten die Befragten in standardisierten Fragebögen sehr stark dazu, die „offizielle“ Darstellung, wonach Hochschulabsolventen in der DDR durch zentrale Vermittlungsbüros bei den Universitäten ihren Arbeitsstellen zugeteilt wurden, zu bestätigen (vgl. Kelle 2008, S. 256f.). In intensiven qualitativen Interviews zeigte sich dann aber, dass der Prozess der Arbeitsvermittlung in der Praxis oftmals ganz anders verlief und die Absolventenvermittlungen vielfach nur eine legitimierende Funktion übernahmen für einen „grauen Arbeitsmarkt“, der im Sozialismus in dieser Form gar nicht existieren sollte. Ein anderes Beispiel kommt aus dem Bereich stationärer Altenpflege. Hier zeigen empirische Untersuchungen, dass die zur Messung von „Kundenzufriedenheit“ in Altenheimen oft eingesetzten Fragebogen-Instrumente zu grob verzerrten Ergebnissen führen (das heißt konkret: tatsächliche Zufriedenheit überschätzen), weil die Befragten sich oft nicht trauen, ihre Unzufriedenheit mit der Pflegesituation zu benennen (Kelle 2007; Kelle, Niggemann, Metje 2008).

Im Kontext von Evaluation und evidenzbasierter Politik lassen sich solche Fehler in Anknüpfung an eine bekannte, zum Sprichwort geronnene historische Anekdote als „Potemkin-Effekt“ bezeichnen: durch elaborierte methodische Werkzeuge wird dann (oft gar nicht beabsichtigt) ein vordergründiger Schein von Veränderung und positiver Entwicklung erzeugt. Solche Probleme werden in einer ausschließlich quantitativen Erhebung häufig nicht sichtbar.

4 Methodenintegration in der Evaluationsforschung

Wie bereits erläutert, wird in der Diskussion um Mixed Methods davon ausgegangen, dass die Kombination quantitativer und qualitativer Erhebungsmethoden das Potential bietet, die Schwächen beider Methodentraditionen durch die Stärken der jeweils anderen zu kompensieren. Die Zielrichtung kann dabei zum einen eine gegenseitige Methodenkritik sein, wenn Methodenprobleme und Validitätsbedrohungen, die bei der qualitativen oder quantitativen Forschung auftauchen, mit Verfahren der jeweils anderen Forschungstradition identifiziert und in vielen Fällen auch behoben werden. Hierbei kann zum anderen eine Ergänzung von Perspektiven erfolgen, weil qualitative und quantitative Methoden den Blick oft auf jeweils unterschiedliche Phänomene oder zumindest unterschiedliche Aspekte desselben Phänomens zu richten helfen (Kelle 2008, S. 54). Diese beiden Funktionen der Methodenkombination, wechselseitige Methodenkritik und Validierung von Ergebnissen einerseits,

Ergänzung von Perspektiven andererseits, widersprechen sich keineswegs, sondern lassen sich oft gemeinsam nutzen – dies möchten wir im letzten Abschnitt unseres Beitrags zuerst methodologisch erläutern und dann anhand eines empirischen Beispiels zeigen.

4.1 *Mixed Methods - Validierungsstrategie oder Perspektivenergänzung?*

In der methodologischen Debatte wird oft der Begriff der „Triangulation“ für eine Kombination von Methoden verwendet (vgl. Kelle 2008, S. 49; Kelle, Erzberger 2001; Erzberger, Kelle 2002). Diese aus der Trigonometrie und Landvermessung entlehnte Metapher wird allerdings von unterschiedlichen Autoren in verschiedener Weise interpretiert. Eine mit den Arbeiten Donald Campbells begonnene Tradition versteht unter Triangulation die Überprüfung und Validierung von Ergebnissen, die mit bestimmten Verfahren der Datenerhebung und –auswertung gewonnen wurden, mit Hilfe anderer Methoden (Campbell, Fiske 1959). Gegen diese Sichtweise ist eingewandt worden, dass qualitative und quantitative Verfahren Prämissen jeweils unterschiedlicher *Theorietraditionen* in den Forschungsprozess einbringen, und deshalb weniger zur gegenseitigen Validierung als zur gegenseitigen Ergänzung geeignet seien (z.B. Fielding & Fielding 1986, S. 33; Flick 1992). Hier zeigen sich die Grenzen des Triangulationsbegriffs ebenso wie seine systematische Vieldeutigkeit: der Begriff 'Position eines Ortes', klar verständlich im Kontext der Landvermessung, ist in der empirischen Sozialforschung nicht genau definiert. Dabei existieren zwei Lesarten der Triangulationsmetapher: Triangulation als kumulative Validierung von Forschungsergebnissen und Triangulation als Ergänzung von Perspektiven, die eine komplexere Erfassung, Beschreibung und Erklärung eines Gegenstandsbereichs ermöglichen. Erfahrungen aus der Forschungspraxis (Kelle 2008, S. 227ff.) machen jedoch deutlich, dass sich das Verhältnis zwischen qualitativen und quantitativen Forschungsergebnissen nicht auf Grund eines einzelnen Modells bestimmen lässt - etwa in dem Sinne, dass Ergebnisse qualitativer und quantitativer Methoden grundsätzlich konvergent sind und deswegen zur gegenseitigen Validierung verwendet werden können oder in dem Sinn, dass sich qualitative und quantitative Ergebnisse unter jeweils verschiedenen Bedingungen zu einem stimmigen Gesamtbild verbinden lassen. Bei einem parallelen Einsatz qualitativer und quantitativer Verfahren in einem gemeinsamen Untersuchungsdesign sind vielmehr drei Ausgänge möglich (ebd., S. 232). In manchen Fällen *konvergieren* qualitative und quantitative Forschungsergebnisse, in anderen Fällen verhalten sie sich *komplementär* zueinander (ergänzen sich also), manchmal sind sie auch *diver-*

gent, d.h. sie widersprechen sich gegenseitig. Insbesondere das Vorliegen von Divergenzen ist für unsere Überlegungen von besonderem Interesse – in der Regel machen solche Widersprüche nämlich Probleme und Grenzen von Methoden, die einer der beiden Traditionen entstammen, deutlich. Hiermit wird also jeweils ein methodisches Problem aufgeworfen, wobei eine Erklärung der Divergenz (die ggf. eine Hinzuziehung weiterer externer Gesichtspunkte, etwa in Form neuer theoretischer Erklärungen, erfordern kann) dann wiederum zu einer Komplementarität der qualitativen und quantitativen Forschungsergebnisse führt. Das bedeutet: Die zusätzlichen Untersuchungen mit Methode 2, deren Resultate den Ergebnissen von Methode 1 widersprechen, können helfen, die Daten von Methode 1 in einem neuen und anderen Licht zu beurteilen und zu interpretieren, so dass beide Befunde aufeinander bezogen werden können und auf diese Weise ein neues und vollständigeres Bild des untersuchten Gegenstandsbereichs ergeben.

Dies lässt sich gut zeigen anhand empirischer Beispiele aus Projekten, in denen Methodenkombination genutzt wurde, um die Begrenzungen, welche die „Monomethodenforschung“ bei der Erfassung und Beschreibung konkreter empirischer Phänomene hat, zu erkennen und auszugleichen. Wir möchten uns dabei im Folgenden auf die Nutzung qualitativer Forschung für die Identifikation von Validitätsbedrohungen und Methodenproblemen quantitativer Methoden konzentrieren, weil eine solche Form der Methodenkombination gerade in quantitativen Evaluationsprojekten bislang nur selten systematisch genutzt wird.

4.2 Qualitative Forschung zur Identifikation von Methodenproblemen und Validitätsbedrohungen quantitativer Forschung

Quantitative Methoden reichen oftmals nicht aus, um alle potentiellen Wirkungen von Interventionsmaßnahmen zu erfassen, weil zur Konstruktion der in der quantitativen Forschung verwendeten Instrumente immer die interessierenden Effekte *vor* einer empirischen Datenerhebung expliziert werden müssen. Somit sind quantitative Methoden nicht geeignet, um unbekannte Phänomene und Handlungsweisen von Akteuren im Feld zu entdecken. Wenn diese Handlungen dann in der oben beschriebenen Weise zu Störvariablen werden, die dem Effekt der unabhängigen Variablen (also der Intervention) entgegen wirken, steigt das Risiko, dass die Nullhypothese (also die Annahme, dass die Maßnahme keinen Erfolg erzielt hat) beibehalten wird und die Hypothese, dass die Maßnahme einen positiven Effekt erzeugt hat, dann zurückgewiesen

werden muss (obwohl der Effekt tatsächlich nur durch den Einfluss von Störvariablen, die in vielen Fällen eliminiert werden könnten, vermindert wird).

Eine wesentliche Stärke qualitativer Verfahren in der empirischen Sozialforschung besteht darin, dass hiermit bislang *unbekannte und nicht in den Vorannahmen des Forschers explizit benannte Phänomene* systematisch in den Blick genommen werden können. Auf diese Weise können etwa die (unbeabsichtigten) erwünschten und unerwünschten Nebeneffekte von Interventionen untersucht werden.

Werden bei der Evaluation eines Modellprojektes wie „Jedem Kind ein Instrument“ nicht nur standardisierte Fragebögen und andere Messinstrumente eingesetzt zur Erhebung vorab definierter Variablen, deren Beeinflussung man sich durch die Intervention erhofft (wie etwa die bessere Befähigung zur Stressbewältigung, die verbesserte auditive Wahrnehmung die musikalischen Präferenzen u.a.m. betreffend), sondern auch Verfahren der qualitativen Datenerhebung (wie offene, teilstandardisierte Interviews), kann hiermit ein breites Spektrum an Effekten exploriert werden, die die Maßnahme aus Sicht der Betroffenen (also der unterrichteten Kinder, ihrer Eltern, den Lehrern und Schulleitungen) hat. Manche unerwünschten Nebeneffekte von sozial- und bildungspolitischen Maßnahmen (etwa ein erhöhtes Mobbingrisiko für hörgeschädigte Kinder in Integrationsklassen) lässt sich nur mit Hilfe von qualitativen Methoden (also etwa durch die Erhebung von Tiefeninterviews oder durch intensive teilnehmende Beobachtung vor Ort) überhaupt beobachten und einer wissenschaftlichen Untersuchung zugänglich machen.

Insbesondere die *Entdeckung unbekannter Störvariablen* durch die Handlungen von Akteuren, die der Erreichung der Programmziele beabsichtigt oder unbeabsichtigt entgegen wirken, ist ohne einen intensiven Kontakt mit dem Forschungsfeld oft gar nicht möglich. Die beteiligten Forscher müssen hier oft vor Ort Beobachtungen machen und Gespräche führen und zum Teil Beziehungen zu den Akteuren im Feld aufbauen, um Informationen über Prozesse zu erhalten, die die Umsetzung von Maßnahmen blockieren, behindern oder verzögern können.

Dies ist vor allem dann der Fall, wenn die untersuchten Interventionen noch nicht in der Sphäre der öffentlichen Meinung mit Rede und Gegenrede debattiert werden (wie dies etwa bei den Maßnahmen des „busing“ der Fall war, wo eine empirische Untersuchung der Gegenstrategien von betroffenen Akteuren nicht ohne weiteres durch eine Inhaltsanalyse von Medienberichten erfolgen kann). Wenn das Untersuchungsfeld etwa eine bestimmte Organisation bildet, in der eine Maßnahme durchgeführt wird, dürfen etwa die „mikropolitischen“ Strategien der beteiligten Akteure nicht vernachlässigt werden – dies ist oft ohne professionelle ethnographische Strategien des Feldzugangs gar nicht möglich.

Schließlich lassen sich qualitative Methoden auch gut dafür einsetzen, um *Quellen unbeobachteter Heterogenität* zu finden. Werden, angeleitet durch eine geschickte Fallkontrastierung nach der Methode des von Glaser und Strauss

vorgeschlagenen „*theoretical sampling*“ (Glaser, Strauss 1967, S. 45 ff.) systematisch unterschiedliche Akteure befragt und verschiedene Situationen und *settings* im Feld vergleichend analysiert, lassen sich auf diese Weise Informationen sammeln über unterschiedliche Bedingungskonstellationen, unter denen bestimmte Maßnahmen besonders gut oder besonders schlecht wirken.

Im Fall des Modellprojekts „Jedem Kind ein Instrument“ würde das bedeuten, dass mit Angehörigen der verschiedenen Akteursgruppen Interviews durchgeführt und in sehr unterschiedlichen Schulen und Unterrichtssituationen Beobachtungen gemacht werden. Auch hier hat ein Mixed Methods Design wiederum besondere Stärken gegenüber einem Monomethodendesign: wenn mit Hilfe quantitativer Untersuchungen und Mittelwertsvergleichen systematisch Schulen oder Schülergruppen, in denen das Programm besonders großen Erfolg und solche Einrichtungen und Schülergruppen, in denen es geringen oder keinen Erfolg bezogen auf die untersuchten outcome-Variablen aufweist, verglichen werden, kann in einem nächsten Untersuchungsschritt mit qualitativen Methoden untersucht werden, *aus welchen Gründen* (das heißt: unter welchen Bedingungen) die Intervention erfolgreich ist oder nicht.

Darüber hinaus können im Rahmen der Evaluationsforschung systematische Messfehler und Methodenartefakte standardisierter Messinstrumente durch ergänzend eingesetzte qualitative Verfahren aufgespürt und auch beseitigt werden. Dabei kann die gesamte Palette der in der Literatur seit langem bekannten Methodenprobleme standardisierter Befragung mit unterschiedlichen qualitativen Verfahren aufgedeckt und bearbeitet werden: Verständnisprobleme der Befragten, Interviewereffekte, Effekte sozialer Erwünschtheit u.a.m. lassen sich auf diese Weise in den Blick nehmen und es können Maßnahmen getroffen werden, um diese Probleme in den Griff zu bekommen.

Validitätsprobleme, die quantitative Befragungen bei der Erhebung der Zufriedenheit von Bewohnern stationärer Einrichtungen der Altenpflege aufweisen (siehe oben) konnten wir bspw. in einer Mixed Methods Studie detailliert beschreiben und darstellen. In dieser Untersuchung wurden qualitative und quantitative Methoden kombiniert, um die Qualität von stationären Pflegedienstleistungen zu untersuchen. Hierzu wurden qualitative Leitfadeninterviews sowie eine Befragung mit einem standardisierten Fragebogen in einer Reihe von Pflegeeinrichtungen mit unterschiedlicher Bettenzahl und Trägerschaft in ganz Deutschland durchgeführt. Ein Teil der standardisierten *face-to-face* Interviews wurde dabei auf Tonträger aufgezeichnet und transkribiert. Erst durch diese Kombination qualitativer und quantitativer Methoden konnte aufgedeckt werden, dass eine standardisierte Befragung bezogen auf Variablen, mit denen die Zufriedenheit der Heimbewohner mit bestimmten Dienstleistungen der Einrichtung gemessen werden sollten, ein aus der Perspektive der untersuchten Heimbewohner hochgradig verzerrtes Bild wiedergab: Zahlreiche Befragte neigen in standardisierten Interviews systematisch dazu, aus Furcht vor Sanktionen Kritik und Unzufriedenheit an der Einrichtung und dem Pflegepersonal zu verschweigen – in qualitativen Intensivinterviews werden solche Probleme aber dann sichtbar und

einer Untersuchung zugänglich (vgl. Kelle, Niggemann 2002; Kelle, Niggemann, Metje 2008).

Im Folgenden möchten wir noch einmal anhand eines eigenen empirischen Beispiels aus der Evaluationsforschung an Hochschulen darstellen, wie die Beschränkungen standardisierter Instrumente, die in Monomethodendesigns nahezu zwangsläufig unentdeckt bleiben, durch eine Kombination qualitativer mit quantitativen Methoden bearbeitbar werden.

4.3 Ein empirisches Beispiel: Die Identifizierung von Itemverständnisproblemen durch qualitative Methoden

Wie wir schon oben ausgeführt haben, stellt in quantitativen Untersuchungen die Verständlichkeit der Items ein grundlegendes Qualitätskriterium dar. Valide Ergebnisse sind nur dann zu erheben, wenn die Items so präzise formuliert sind, dass sie von allen Befragten auf die gleiche Weise verstanden werden und wenn ausreichendes *Wissen* zur angemessenen Beantwortung vorhanden ist (Fowler 1995, 4). Validitätsbedrohungen durch divergierende kognitive Prozesse der Befragten können allerdings nur dann identifiziert werden, wenn qualitative Methoden ergänzend zu den quantitativen Erhebungen eingesetzt werden. Das Mittel der Wahl stellt hier die Technik des kognitiven Interviews dar (vgl. Willis 2005, 42ff.), bei dem die Befragten während des Ausfüllens eines Fragebogens entweder gebeten werden, alle Gedanken laut auszusprechen (Think-Aloud-Technik), die Frage in eigenen Worten wiederzugeben, die Verlässlichkeit ihrer Antworten zu bewerten oder komplizierte Begriffe zu erklären (in der Methodenliteratur als Probing bezeichnet). Offenbaren sich dabei Verständnisdivergenzen zwischen den Befragten, müssen die entsprechenden Items überarbeitet werden.

Am Beispiel eines Fragebogenitems zur Lehrveranstaltungsevaluation soll gezeigt werden, wie die Itemqualität durch qualitative Methoden überprüft und verbessert werden kann. Eine Aussage aus den FEVOR- und FESEM-Fragebögen (Staufenbiel 2000), die sich in ähnlicher Weise auch in anderen gebräuchlichen Erhebungsinstrumenten zur Lehrevaluation findet, lautet: „Die Vorlesung/ Das Seminar gibt einen guten Überblick über das Themengebiet.“

In den Ausführungen der befragten Studierenden wird deutlich, dass sie das Item nur angemessen beantworten können, wenn sie selbst schon einen Überblick über die Breite und Komplexität des Themengebietes gewonnen haben. Somit ist das Item besonders für Studierende in den Anfangssemestern kaum verwendbar und auch in höheren Semestern wird der Horizont der Studierenden häufig nicht so weit über die Thematik der Veranstaltung hinausge-

hen, dass eine valide Einschätzung vorgenommen werden kann. In den Interviewpassagen finden sich dazu u.a. folgende Aussagen²:

„je, nachdem, welchen Dozenten sie haben, ist es auch sehr einseitig ausgerichtet. Zum Beispiel [...] der Herr B. [...], der ist halt ein großer Anhänger der Verhaltenstherapie und die Tiefenpsychologie wird so ganz ausgeblendet. Aber am Anfang merken sie das gar nicht. [...] durch Praktika oder dadurch, dass man etwas mehr Zugang zur Materie kriegt, merkt man, dass es vielleicht doch ein bisschen einseitig ist.“ (Interview Nr. 7, 4.10.2007, Abs. 30)

„Ja, es ist schwierig. Es ist halt das erste Mal, dass ich mich mit dem Thema befasse. Jetzt würde ich natürlich sagen, es gibt einen sehr guten Überblick, aber, unter Umständen gibt's ja auch was, was wir nicht behandeln und was ich dann ja auch gar nicht mitkriege.“ (Interview Nr. 18, 7.1.2008)

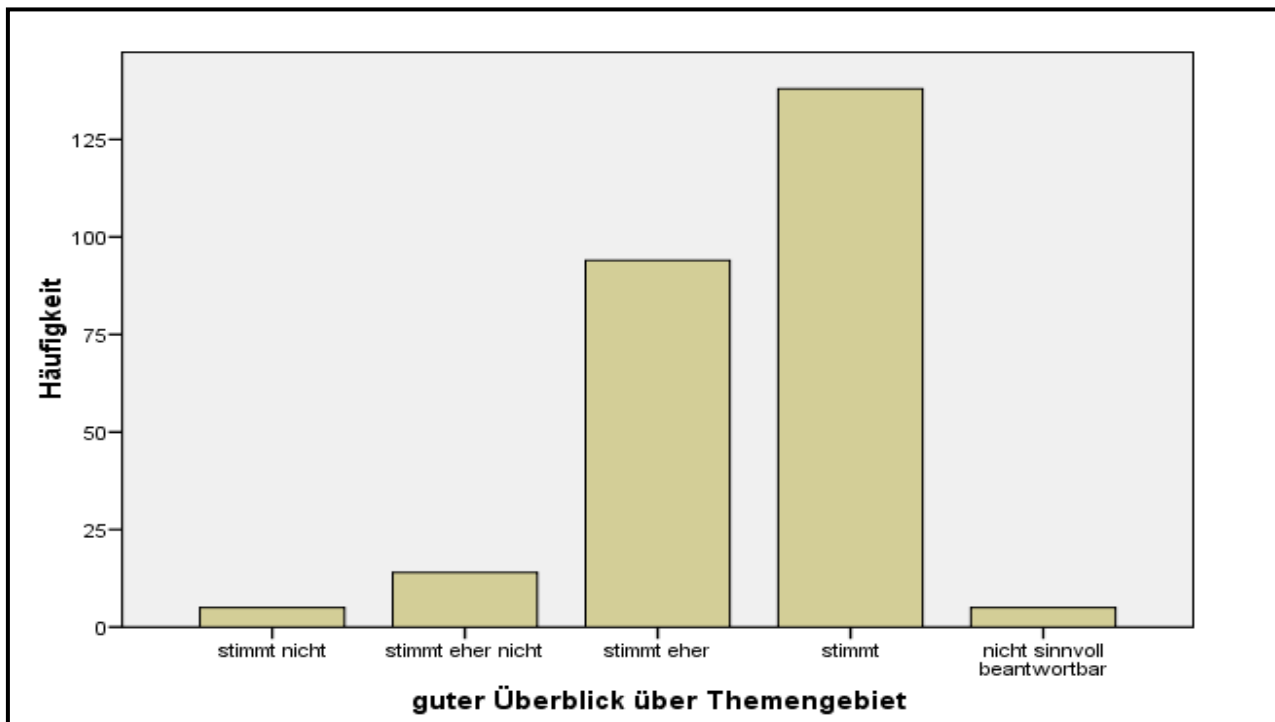
„Man kann hoffen, dass die Dozenten das auswählen, was es da so alles gibt. Weil man ja selber da noch nicht so den Überblick hat. Und darauf hoffen, dass die eben das Relevante auswählen.“ (Interview Nr. 6, 24.9.2007)

„Also, das ist schwer zu sagen, ob es jetzt ein guter Überblick war. Bis jetzt kennen wir ja noch nicht so viel Anderes. Also von daher „stimmt eher“.“ (Interview Nr. 5, 23.8.2007)

„Kann ich bei den meisten Seminaren, glaube ich, nicht abschätzen. Weil mir da das große Ganze fehlt. Für gewöhnlich lernt man den Großteil ja erst im Seminar. Gut, da habe ich jetzt einfach mal Vertrauen. Insgesamt meine ich, es passt. [...] Deshalb "stimmt eher".“ (Interview Nr. 15, 17.12.2007, Abs. 53-58)

Diese Unsicherheiten der Studierenden werden in den quantitativen Analysen kaum sichtbar. Selbst wenn die Studierenden sich nicht kompetent fühlen, das Item zu beantworten, wählen sie vorwiegend eine zustimmende Antwortalternative und geben dem Fragebogenentwickler nur selten den Hinweis, das Item sei „nicht sinnvoll beantwortbar“. Schauen wir uns dazu die Verteilung an, die sich in einer Evaluation von neun Veranstaltungen an der Universität Marburg ergab (vgl. Grafik 3).

2 Metje, 2009, 232ff.



Grafik 3

Lediglich 8% der befragten 357 Studierenden wählte eine ablehnende Antwortalternative und 4% hielt die Aussage für „nicht sinnvoll beantwortbar“. Verlässt sich der Evaluator nun nur auf die quantitativen Ergebnisse, werden die in den qualitativen Interviews offenbar gewordenen Validitätsbedrohungen nicht evident. Nur dann, wenn quantitative und qualitative Daten systematisch aufeinander bezogen werden, können Probleme dieser Art identifiziert werden. Das obengenannte Item sollte letztendlich gänzlich aus dem Fragebogen entfernt werden, weil die Studierenden offenbar nicht in der Lage sind, diesen Qualitätsaspekt von Lehre adäquat einzuschätzen.

5 Zusammenfassung und Ausblick

Mit Evaluationen werden Kausalbeziehungen untersucht, indem in der Regel ausschließlich auf der Basis quantitativer und standardisierter Erhebungsinstrumente die vorab definierten Wirkungen spezifischer Maßnahmen erhoben werden. In unserem Beitrag haben wir gezeigt, dass eine ausschließliche Konzentration des methodischen Instrumentariums auf quantitative Erhebungsinstrumente mit verschiedenen Validitätsbedrohungen einhergehen kann. So können entweder Veränderungen fälschlicherweise konstatiert oder auch übersehen werden oder es werden bestimmte nicht antizipierte Kausalzusammen-

hänge übersehen, weil die Handlungsziele von Akteuren nicht ohne weitere Exploration vorhersehbar und die Handlungsbedingungen variabel sind und darüber hinaus Handlungsregeln, denen die Akteure unterliegen, nur für eng begrenzte, lokale Handlungsfelder gelten oder unbekannten soziokulturellen Veränderungsprozessen unterliegen. Ein theoretisch angenommenes lineares Zusammenhangsmodell zwischen Intervention und vorher definiertem Outcome scheitert damit letztlich in der Realität an der Flexibilität und „Kreativität des Handelns“ (Joas 2002) der Akteure im Feld, die je eigene Interessen und Ziele verfolgen können. Entwickler quantitativer Erhebungsinstrumente können oftmals diese individuellen Wirkungspfade und unbeabsichtigten Nebeneffekte von Interventionen nicht antizipieren und damit auch nicht in Form von Items operationalisieren. An diesem Punkt werden die Vorteile einer Methodenintegration in der Evaluationsforschung offensichtlich: mit qualitativen Methoden können Informationen über die Handlungsziele und Handlungsweisen von Akteuren gesammelt und unvorhersehbare Outcomevariablen definiert, operationalisiert und präzisiert werden. Sollen die Ausmaße solcher Effekte quantifiziert werden, bietet sich dem Forscher jetzt die Möglichkeit, entsprechende Items in den Fragebogen aufzunehmen.

Eine weitere Validitätsbedrohung ausschließlich quantitativer Erhebungen stellen systematische Messfehler und Methodenartefakte dar. Sollen Fragebögen zur Evaluation spezifischer Interventionen entwickelt werden, ist eine Einbeziehung qualitativer Verfahren unverzichtbar, um divergierende Iteminterpretationen der Befragten zu entdecken und uneindeutige Itemformulierungen zu präzisieren. Auch Items, die von den Befragten nicht adäquat beantwortet werden können, weil ihnen das entsprechende Wissen fehlt, können mit kognitiven Interviewtechniken identifiziert und dann entsprechend überarbeitet werden. Allerdings haben diese Validitätsproblematiken unter deutschen Evaluationsforscher(inne)n bisher nur wenig Beachtung gefunden. Bisher ist es hierzulande, anders als in Amerika, wo eine Debatte über kognitive Prozesse bei der Beantwortung von Fragebögen schon seit den frühen 1980er Jahren geführt wird, nicht üblich, die Qualität quantitativer Instrumente durch qualitative Daten abzusichern. Unsere Ausführungen sollen deshalb ein Beitrag dazu sein, das Bewusstsein für die Potentiale und die Notwendigkeit einer Methodenintegration zu schärfen. Wie deutlich geworden sein sollte, bietet die Evaluationsforschung zahlreiche Anknüpfungspunkte für die Entwicklung von Konzepten und Methoden der Mixed Methods Forschung – eine stärkere Berücksichtigung der aktuellen Debatten um Methodenintegration würde die Methodenentwicklung in diesem Feld also erheblich voranbringen.

Literatur:

- ARNOLD, A. (1972/1662): Die Logik oder die Kunst des Denkens. Darmstadt: Wissenschaftliche Buchgesellschaft.
- BABAD, E.; KATZ, Y. (1991): Wishful Thinking – Against All Odds. In: *Journal of Applied Social Psychology*, 21, pp. 1921-1938.
- BENECKEN, J.; SPINDLER, S. (2004): Zur psychosozialen Situation stotternder Schulkinder in Allgemeinschulen. In: *Die Sprachheilarbeit*, 49 (2), S. 61 – 70.
- BLALOCK, H.M. (1985): Causal Models in the Social Sciences. New York: Aldine.
- BRYMAN, A. (2007): Barriers to integrating quantitative and qualitative research. In: *Journal of Mixed Methods Research*, 1, pp. 1-18.
- BUNDESMINISTERIUM FÜR BILDUNG UND FORSCHUNG (2001): TIMSS- Impulse für Schule und Unterricht: Forschungsbefunde, Reform-initiativen, Praxisberichte und Video-Dokumente. Bonn: BMBF publik.
- CAMPBELL, D.T.; STANLEY, J.C. (1963): Experimental and Quasi-Experimental Designs for Research. Dallas: Houghton Mifflin.
- CAMPBELL, D.T.; FISKE, D.W. (1959): Convergent and Discriminant Validation by the Multitrait-Multimix Matrix. In: *Psychological Bulletin*, 56, pp. 81-105.
- CHALMERS, A.F. (2007): Wege der Wissenschaft: Einführung in die Wissenschaftstheorie. Berlin: Springer.
- CICOUREL, A. V. (1974): Methode und Messung in der Soziologie. Frankfurt am Main: Suhrkamp.
- COLEMAN, J.S. (1976): Liberty and equality in school desegregation. In: *Social Policy*, 6, pp. 9-13.
- COLLINGWOOD, R.G. (1937/1938): On the So-called Idea of Causation. In: *Proceedings of the Aristotelian Society*, 38, pp. 85-112.
- CRESWELL, J.W.; PLANO CLARK, V.L.; GUTMANN, M.L.; HANSON, W.E. (2003): Advanced mixed methods research designs. In: Tashakkori, A.; Teddlie, C. (Eds.): *Handbook of Mixed Methods in Social and Behavioral Sciences*. Thousand Oaks, Calif.: Sage, pp. 209-240).
- CRESWELL, J.W. (2009): Research design: qualitative, quantitative and mixed methods approaches. Thousand Oaks, Calif.: Sage.

- ERZBERGER, C.; KELLE, U. (2002): Making Inferences in Mixed Methods: The Rules of Integration. In: Tashakkori, A. & Teddlie, C. (Eds.). Handbook of mixed methods for the social and behavioural sciences. Thousand Oaks, Calif.: Sage, pp. 457 – 490.
- FESTINGER, L. (1957): A theory of cognitive dissonance. Stanford, Cal.: Stanford University Press.
- FESTINGER, L.; RIECKEN, H.W.; SCHACHTER, S. (1956): When Prophecy Fails. Minneapolis, Mi.: University of Minnesota Press.
- FIELDING, N.G.; FIELDING, J.L. (1986): Linking Data. (*Qualitative Research Methods Vol. 4*). London: Sage.
- FLICK, U. (1992): Triangulation Revisited: Strategy of Validation or Alternative? In: *Journal for the Theory of Social Behaviour*, 22, pp. 175-197.
- FOWLER, F.J. (1995): Improving survey questions: design and evaluation. Thousand Oaks: Sage.
- GLASER, B.; STRAUSS, A. (1967): The Discovery of Grounded Theory. Strategies for qualitative Research. New York: Aldine.
- GREENE, J.C. (2007): Mixed Methods in Social inquiry. San Francisco, Calif.: Jossey Bass.
- GUBA, E.G.; LINCOLN, Y.S. (1988): Do inquiry paradigms imply inquiry methodologies? In: Fetterman, D.M. (Ed.) Qualitative approaches to evaluation in education: The silent scientific revolution. New York: Praeger, pp. 88-115.
- HAGE, J.; MEEKER, B.F. (1988): Social Causality. Boston: Unwin Hyman.
- HEINZ, W.R.; MARSHALL, V.W. (2003): Social Dynamics of the Life Course: Transitions, Institutions, and Interrelations. New York: Aldine de Gruyter.
- JAHODA, M.; LAZARFELD, P.F.; ZEISEL, H. (1933/1982): Die Arbeitslosen von Marienthal. Frankfurt: Suhrkamp.
- JOAS, H. (2002): Die Kreativität des Handelns. Frankfurt/M.: Suhrkamp.
- JOHNSON, B.; TURNER, L.A. (2003): Data Collection Strategies in Mixed Methods Research. In: Tashakkori, A.; Teddlie, C. (Eds.): Handbook of Mixed Methods in Social and Behavioral Sciences. Thousand Oaks, Calif.: Sage, pp. 297-319.

- KAHNEMAN, D. (1982/2008): Judgment under uncertainty: heuristics and biases. Cambridge: Cambridge University Press.
- KELLE, U. (2006): Combining Qualitative and Quantitative Methods in Research Practice – Purposes and Advantages. In: Gürtler, L.; Huber, G.L. (Eds.). Special Guest Issue on Mixed Methods. Qualitative Research in Psychology. Vol. 3 (4), pp. 293-311.
- KELLE, U. (2007): “Kundenorientierung” in der Altenpflege? Potemkinsche Dörfer sozialpolitischen Qualitätsmanagements. In: *PROKLA 146, Zeitschrift für kritische Sozialwissenschaft*, 37 (1), S. 113-128.
- KELLE, U. (2008): Die Integration qualitativer und quantitativer Methoden in der empirischen Sozialforschung. Theoretische Grundlagen und methodologische Konzepte. Wiesbaden: VS Verlag für Sozialwissenschaften.
- KELLE, U.; ERZBERGER, C. (2001): Die Integration qualitativer und quantitativer Forschungsergebnisse. In: KLUGE, S.; KELLE, U. (Hg.): Methodeninnovation in der Lebenslaufforschung. Integration qualitativer und quantitativer Verfahren in der Lebenslauf- und Biographieforschung. Weinheim; München: Juventa, S. 89- 133.
- KELLE, U.; NIGGEMANN, C. (2002): „Wo ich doch schon einmal vor zwei Jahren verhört worden bin...“: Methodische Probleme bei der Befragung von Heimbewohnern. In: Motel-Klingebiel, A.; Kelle, U. (Hg.): Perspektiven der empirischen Alterssoziologie. Opladen: Leske und Budrich, S. 99 – 132.
- KELLE, U.; NIGGEMANN, C.; METJE, B. (2008): Datenerhebung in totalen Institutionen als Forschungsgegenstand einer kritischen gerontologischen Sozialforschung. In: Amann, A.; Kolland, F. (Hg.): *Das erzwungene Paradies des Alters?*, S. 163-193.
- KLUGE, S.; KELLE, U. (2001) (Hg.): Methodeninnovation in der Lebenslaufforschung. Integration qualitativer und quantitativer Verfahren in der Lebenslauf- und Biographieforschung. (Statuspassagen und Lebenslauf, Band 4). Weinheim; München: Juventa.
- KRANEFELD, U. (2009): Perspektivwechsel: Den musikalischen Denkwegen der Schülerinnen und Schülern folgen. In: Greuel, T., Kranefeld, U.; Szczepaniak, E. (Hg.): Jedem Kind (s)ein Instrument. Die Musikschule in der Grundschule. (Musik im Diskurs; 23). Aachen: Shaker.
- LAMNEK, S. (1995): Qualitative Sozialforschung. Bd. 1: Methodologie. Weinheim: Beltz, Psychologie-Verlags-Union.

LANDESHOCHSCHULGESETZE:

http://www.hof.uni-halle.de/steuerung/lhg_uebersicht.htm, letzter Zugriff am 6.4.2010)

- LAZARSFELD, P.F. (1955): Interpretation of Statistical Relations as Research Operation. In: Lazarsfeld, P.D.; Rosenberg, M. (Eds.): *The Language of Social Research*. New York: John Wiley and Sons, pp. 115-125.
- LINCOLN, Y.S.; GUBA, E.G. (2000): Paradigmatic controversies, contradictions, and emerging confluences. In: Denzin, N.K.; Lincoln, Y.S. (Eds.): *Handbook of Qualitative Research*. Thousand Oaks: Sage, pp. 163-188.
- METJE, B. (2009): *Evaluation universitärer Lehrveranstaltungen. Standpunkte von Studierenden und Systemdefizite im Fokus*. Marburg: Tectum.
- ONWUEGBUZIE, A.J.; WITCHER, A.E.; COLLINS, K.M.T.; FILER, J.D.; WIEDMAIER, C.D.; MOORE, C.W. (2007): Students' Perceptions of Characteristics of Effective College Teachers: A Validity Study of a Teaching Evaluation Form Using a Mixed-Methods Analysis. In: *American Educational Research Journal*, 44 (1), pp. 113-160.
- PRENZEL, M.; ARTELT, C.; BAUMERT, J.; BLUM, W.; HAMMANN, M.; KLIEME, E.; PEKRUN, R. (2007) (Hg.): *Pisa 2006: Die Ergebnisse der dritten internationalen Vergleichsstudie*. Münster; New York; München; Berlin: Waxmann.
- PRÜFER, P.; REXROTH, M. (2005): *Kognitive Interviews. ZUMA How-to-Reihe*, Nr. 15. [http://www.gesis.org/Publikationen/Berichte/ZUMA_How_to/Dokument_e/pdf/How_to15PP_MR.pdf, 14.3.2007]
- ROETHLISBERGER, F.J.; DICKSON, W.J. (1939): *Management and the worker*. Cambridge, Ma.: Harvard University Press.
- SCHÜTZ, A. (1971): *Gesammelte Aufsätze. Bd. 1: Das Problem der sozialen Wirklichkeit*. Den Haag: Nijhoff.
- SEIPEL, C.; RIEKER, P. (2003): *Integrative Sozialforschung. Konzepte und Methoden der qualitativen und quantitativen empirischen Forschung*. Weinheim; München: Juventa.
- SIMON, H.A. (1954): Spurious Correlation: A Causal Interpretation. In: *Journal of the American Statistical Association*, 49, pp. 467-479.
- SIMPSON, C.H. (1951): The Interpretation of Interaction in Contingency Tables. In: *Journal of the Royal Statistical Society*, 13, pp. 238-241.

- STAUFENBIEL, T. (2000): Fragebogen zur Evaluation von universitären Lehrveranstaltungen durch Studierende und Lehrende. In: *Diagnostica*, 46 (4), S. 169-181.
- STRAUSS, A.; CORBIN, J. (1990): Basics of Qualitative Research. Grounded Theory Procedures and Techniques. Newbury Park, Calif.: Sage.
- TASHAKKORI, A.; TEDDLIE, C. (2003) (Eds.): Handbook of Mixed Methods in Social and Behavioral Sciences. Thousand Oaks, Calif.: Sage.
- WHITE, H. (2006): Impact Evaluation: The Experience of the Independent Evaluation Group of the World Bank. Washington D.C.: World Bank.
- WILLIS, G. B. (2005): Cognitive interviewing: a tool for improving questionnaire design. Thousand Oaks; London; New Delhi: Sage.
- ZIMBARDO, P.G. (1969): The human choice: Individuation, reason and order versus deindividuation, impulse and chaos. In: Arnold, W.T.; Levine, D. (Eds.): Nebraska Symposium on Motivation, 17, pp. 237-307. Lincoln, Ne.: University of Nebraska Press.

Schulentwicklung mit Chor- und Bläserklassen¹

Eine qualitative Fallstudie am „Evangelischen Gymnasium am Dom zu Brandenburg“

1 Einführung

Das mittelmäßige Abschneiden deutscher Schülerinnen und Schüler im Rahmen der internationalen Schulvergleichsstudien wie PISA oder TIMSS hat grundlegende Debatten zur Qualität des deutschen Bildungssystems angestoßen. Die ernüchternden Befunde wurden vielfältig diskutiert und haben einschneidende Veränderungsmaßnahmen in Bezug auf Schul- und Unterrichtsentwicklung ausgelöst. Der Ausbau von Ganztagschulen, die rege Neugründung im Bereich der Schulen in freier, privater oder konfessioneller Trägerschaft sowie die zunehmende Tendenz individueller Profilbildungen an deutschen Schulen sind als positive Ergebnisse dieser Entwicklung zu werten. Weiterhin ist die Diskussion um Bildungsstandards und die damit einhergehende Erstellung von Kompetenzmodellen in den verschiedenen Unterrichtsfächern in dieser Entwicklung zu verorten (Klieme, Avenarius & Blum et al. 2003; Criblez, Oelkers, Reusser, Berner, Halbheer & Huber 2009).

Bei diesen Bestrebungen standen zunächst die Kernfächer im Fokus. Dies erscheint einleuchtend, weil Kompetenzen in Schreiben, Lesen und Rechnen die grundlegende Voraussetzung zur Teilhabe am gesellschaftlichen und politischen Leben bilden. Seit einiger Zeit werden jedoch in zunehmendem Maße die sogenannten „kleinen Fächer“ wie Musik (*siehe auch:* Knigge & Lehmann-Wermser 2008), Kunst, Darstellendes Spiel oder auch Sport (*siehe auch:* Laging, Hildenbrandt-Stramann & Teubner, n. d.) in den Blick genommen. Nicht selten wird diesen Unterrichtsfächern mit einer hohen Erwartungshaltung entgegen getreten: seien es die sogenannten *soft skills*, die sich in künstlerisch-ästhetischen oder gruppendynamischen Schulfächern besser for-

1 Dieser Beitrag ist vom Vorstand des Arbeitskreises Musikpädagogische Forschung mit dem Forschungspreis für das Jahr 2009 ausgezeichnet worden.

men ließen, sei es der vermutete Beitrag zu einer Verbesserung von Schul- und Unterrichtsklima oder die Möglichkeit, mit fachbezogenen Ergebnissen und Lehr-Lernsettings die Schulen nach außen zu öffnen und damit besser im lokalen Umfeld zu verankern. Diese Erwartungen haben oftmals den Blick darauf verstellt, was die genannten Fächer in ihrem eigentlichen Verständnis an fachgebundenem Potenzial bergen. Ausgehend von diesen bildungspolitischen Tendenzen schien es wichtig, das Potenzial der künstlerischen Fächer, insbesondere des Schulfaches Musik an Ganztagschulen umfassend zu betrachten.

1.1 Forschungsrahmen und Zielstellung

Der vorliegende Tagungsbeitrag ist eingebettet in die begleitende Forschung zur Ganztagschulentwicklung in Deutschland. Er basiert auf grundlegenden Erkenntnissen, die im Rahmen der „Studie zur musisch-kulturellen Bildung an Ganztagschulen (MUKUS)“ gewonnen werden konnten (Lehmann-Wermser, Naacke & Nonte, *in Druck*). Während die MUKUS-Studie breit über den musisch-kulturellen Fachbereich an insgesamt 29 Ganztagschulen in drei Bundesländern angelegt war, konzentriert sich das vorliegende Forschungsvorhaben auf die Untersuchung einer weiteren Einzelschule. Gegenstand der Arbeit ist die qualitativ empirische Erforschung einer Schule in konfessioneller Trägerschaft im Bundesland Brandenburg. Das „Evangelische Gymnasium am Dom zu Brandenburg“² stand zum Zeitpunkt der Untersuchung am Anfang einer Schulentwicklungslaufbahn und hat sich im Zuge dessen freiwillig für eine Evaluation des bisherigen Standes zur Verfügung gestellt.

Zentrales Anliegen der Arbeit war es gerade nicht, Effizienz und Leistung einzelner schulischer Akteure oder etwa eines schulischen Systems zu einem Zeitpunkt X zu messen. Vielmehr sollten Wechselwirkungen, Veränderungen und dynamische Interaktionen sowie hemmende oder fördernde Faktoren, denen sich eine Schule im Zusammenhang mit Schulentwicklungsprozessen gegenüber sieht, rekonstruiert werden.

2 In den folgenden Ausführungen werden die Kurzformen „Domgymnasium“ oder „Evangelisches Domgymnasium“ für die Schulbezeichnung bevorzugt, wie sie auch von den schulischen Akteuren verwendet werden

1.2 Das „Evangelische Gymnasium am Dom zu Brandenburg“

Die untersuchte Einzelschule ist ein Gymnasium in kirchlicher Trägerschaft³ in Brandenburg an der Havel und verfolgt einen mehrjährigen Entwicklungsprozess, der den schulischen Ausbau mit zwei Klassen pro Jahrgang pro Folgeschuljahr vorsieht. Da im Bundesland Brandenburg außerdem die sechsjährige Grundschule obligatorisch ist, waren zum Zeitpunkt der Datenerhebung im Mai 2008 lediglich die Jahrgänge sieben und acht mit je zwei Klassen vertreten. Die Schule beruft sich nicht nur auf ihr evangelisches Profil, sondern sie ist zudem eine Ganztagschule in gebundener Form, die für Schülerinnen und Schüler der Jahrgänge sieben und acht die Teilnahme an der Chor- oder an der Bläserklasse verpflichtend zur Auswahl stellt.

Auch wenn die Blickrichtung vornehmlich auf den Musikbereich und die implementierten musikalischen Praxen gerichtet ist, ist zusätzlich dazu eine weite Perspektive eingenommen worden. Diese weite Perspektive ermöglicht die Untersuchung von Schulentwicklungsprozessen im Zusammenhang mit der Ausgestaltung einer spezifischen Schulkultur: Diesbezüglich setzt das Domgymnasium drei Schwerpunkte, die hier lediglich kurz erwähnt werden sollen. Neben dem Konzept der *musikalischen Bildung* (insbesondere aufgrund der Durchführung von Chor- und Bläserklassen) fließen in die Betrachtungen die Besonderheiten eines *evangelischen Schulprofils* sowie Grundsätze der *Ganztagschulkonzeption* ein. Diese drei Bereiche stützen die Entwicklung einer veränderten Lernkultur, was am Beispiel des Lernertrags auf Schülerseite im Zusammenhang mit Chor- und Bläserklassen dargelegt werden soll. Diese bildeten stets den Ausgangs- und Zielpunkt der Überlegungen.

Im vorliegenden Beitrag soll auf folgende Bereiche eingegangen werden: Ausgehend von einer knappen Darlegung des gegenwärtigen Schulentwicklungsdiskurses werden organisations- und systemtheoretische Grundannahmen sowie forschungsmethodologische Orientierungen vorgestellt. Darauf aufbauend werden zwei Ergebniskomplexe präsentiert. Einerseits das aus dem Forschungsprozess hervorgegangene allgemeine Schulmodell und andererseits eine ausschnitthafte Falldarstellung zum Domgymnasium am Beispiel des dreidimensionalen Lernertrags auf Schülerseite. Der Beitrag mündet in die Formulierung eines allgemeinen Merksatzes zur Schulentwicklungstheorie und einer schulspezifischen Erläuterung für das Domgymnasium.

3 Das Domgymnasium ist eine Schule in der Trägerschaft der Schulstiftung der Evangelischen Kirche Berlin-Brandenburg-schlesische Oberlausitz.

2 Schulentwicklungsdiskurs im 21. Jahrhundert

2.1 Hintergrund

Schulentwicklung als ein historisch bedeutsames Anliegen wird seit Jahrhunderten vielfältig thematisiert. Den Ausgangspunkt bildeten die Humboldt'schen Schulreformen im Zusammenhang mit der Durchsetzung staatlicher Reformmaßnahmen (Rahm 2005). Bereits in den 90er Jahren des 20. Jahrhunderts wird die Schule der Zukunft verstanden als „...lernende Organisation, die eine eigene Schulkultur entwickelt und eine Profilbildung vornimmt“ (Rahm 2005, S. 30). Weiterhin werden Anforderungen an die Schule der 90er Jahre formuliert:

„Die Schule möge Mut zu mehr Autonomie aufbringen, sie möge Eigenständigkeit beweisen und ihre Entscheidungs- und Handlungsspielräume vergrößern.“ (Rahm 2005, S. 31)

Die Aufforderung zu mehr schulischer Autonomie und Eigenständigkeit eröffnet den Blick auf organisations- und systemtheoretische Perspektiven, deren Einfluss am Beginn des 21. Jahrhunderts auf Schulentwicklungsprozesse nicht mehr wegzudenken ist. Demgemäß wird Schulentwicklung als Zusammenspiel von Organisations-, Personal- sowie Unterrichtsentwicklung verstanden.

2.2 Modelle zur Erfassung von Schulentwicklungsprozessen

Darauf aufbauend haben sich verschiedene Modelle zur Erfassung von schulischen Veränderungsprozessen etabliert. Zwei sollen beispielhaft vorgestellt werden. Zum einen markiert das Drei-Wege-Modell nach Kempfert & Rolff (1999) unterschiedliche schulpolitische Ebenen von Entwicklungsprozessen, zum anderen werden mit dem Mehrebenenmodell nach Wiater (2002) Entwicklungsverläufe innerhalb einer Einzelschule deutlich.

2.2.1 Das Drei-Wege-Modell der Schulentwicklung (Kempfert & Rolff 1999)

Kempfert & Rolff (1999) beschreiben ein Drei-Wege-Modell der Schulentwicklung. Während die *intentionale Schulentwicklung* (auch Schulentwicklung erster Ordnung) die bewusste systematische Weiterentwicklung der Einzelschule umfasst, wird im Zuge der *institutionellen Schulentwicklung* (zweiter

Ordnung) die Etablierung „Lernender Schulen“ thematisiert. Dies sind Schulen, die in der Lage sind, sich selbst zu organisieren, zu reflektieren und zu steuern. Die *komplexe Schulentwicklung* (dritter Ordnung) wiederum beschreibt Abläufe und Verfahren zur Steuerung des Gesamtzusammenhangs von sich selbst entwickelnden Einzelschulen, sowie von Unterstützungssystemen, von Koordination und Evaluation der Gesamtentwicklung (Kempfert & Rolff 1999, S. 38-39).

2.2.2 Mehrebenenmodell der Schulentwicklung nach Wiater (2002)

Demgegenüber unterscheidet Wiater (2002) *Makrotheorien* zur Organisationsstruktur, *Mesotheorien* zur schulischen Realität (zu Schulkultur und Schulleben) sowie *Mikrotheorien* im Sinne von Handlungs- und Interaktionstheorien auf Unterrichtsebene.

*„Während auf der Unterrichtsebene Maßnahmen zur Verbesserung von Lernerfolgen im Klassenzimmer unternommen werden können, ist auf der Schulebene strukturierte Schulentwicklung durchzuführen; demgegenüber muss das System auf der Makroebene eine Balance zwischen zentraler Steuerung und Kontrolle einzelschulisch zu verantwortender Qualitätsentwicklung finden.“
(Kipper 2002, Fend 2000; zitiert nach Rahm 2005, S. 146)*

Indem in diesem Modell bezogen auf die Einzelschule unterschiedliche Handlungsfelder ineinander greifen, wird zudem die systemische Perspektive deutlich, die Schulentwicklung als Zusammenwirken von Organisations-, Personal- sowie Unterrichtsentwicklung versteht.

2.2.3 Positionierung der Untersuchung am Domgymnasium

Wie am Beispiel des Domgymnasiums präsentiert werden kann, verlaufen die Schulentwicklungsmodelle auf den benannten drei Ebenen nach Kempfert & Rolff (1999) weder losgelöst voneinander noch in klar abgrenzbaren Phasen. Sie gehen ineinander über und bedingen sich gegenseitig. Mit der Einbettung in die Debatte um Ausbau und Weiterentwicklung von Ganztagsschulen in Deutschland, können Eckpunkte einer *komplexen Schulentwicklung* (Schulentwicklung dritter Ordnung) markiert werden. In diesem Zusammenhang ermöglicht die Vorgabe von Konzepten zu unterschiedlichen Ganztagsformen, ebenso wie die Abrufung von Geldern im Rahmen des Investitionsprogramms „Zukunft, Bildung und Betreuung (IZBB)“ der Einzelschule einen Zugriff auf

verschiedene finanzielle und organisatorische Unterstützungssysteme sowie pädagogische Maßnahmen. Weiterhin können am Domgymnasium auch Schulentwicklungsprozesse erster und zweiter Ordnung aufgespürt werden: Die Schule ist aus einem Elternverein hervorgegangen, welcher sich zunächst zur Gründung einer sechsjährigen Grundschule zusammenfand. In den folgenden Jahren ist die Idee zur Gründung eines evangelischen Gymnasiums als weiterführende Option (zumindest für einen Teil der Schulkinder) im Sinne der zugrunde liegenden und gemeinsam geteilten evangelisch-religiösen Haltung verfolgt worden. Sowohl die Schulgründung als auch die damit verbundenen organisatorischen, konzeptionellen, pädagogischen und didaktischen Überlegungen sind Teil einer bewussten systematischen Weiterentwicklung der Einzelschule (*intentionale Schulentwicklung*). Nicht nur anhand der Teilnahme des Domgymnasiums an einem Forschungsvorhaben, auch an schulinternen Maßnahmen (wie bspw. der konzeptionellen Weiterentwicklung von Lernsettings) wird die Bereitschaft des sozialen Systems Schule deutlich, sich als „Lernende Schule“⁴ zu etablieren und somit Schulentwicklungsprozesse zweiter Ordnung zu durchlaufen (*institutionelle Schulentwicklung*).

In Bezug auf das Mehrebenenmodell nach Wiater (2002) wird Folgendes erkennbar: Wenn auf der Unterrichtsebene die Lernfortschritte der Schülerinnen und Schüler im Mittelpunkt stehen, so geht damit die (Weiter-)Entwicklung von Lernarrangements einher, verbunden mit einer entsprechenden Professionalisierung der Lehrkräfte (Personalebene). Auf der Ebene der Organisation wiederum sollen die Potenziale der Organisationsmitglieder mobilisiert werden mit dem Ziel einer Leistungsoptimierung der Organisation. Dies kann mit der Formulierung von gemeinsam getragenen Schulleitbildern und Schulprogrammen erreicht werden. Aber auch die Arbeit im Team sowie die Einrichtung von Steuergruppen unterstützen das gemeinsame Lernen und Kooperieren. Schließlich müssen auf der Organisationsebene Fragen zur Bewirtschaftung der Einzelschule sowie zur konkreten Umsetzung organisationsbezogener Ziele und Vorgaben gestellt und beantwortet werden.

4 Auf die Besonderheiten im Zusammenhang mit einem Verständnis von Schulen als „Lernende Organisationen“ wird im vorliegenden Beitrag nicht eingegangen. Dies kann in Naacke (2010) nachgelesen werden.

2.2.4 Zusammenfassung

Die Theoriebildung zur Schulentwicklung findet ihren Ausgangspunkt bei der Erforschung und Entwicklung pädagogischer Praxis und wirkt auf die Praxis zurück, indem sie die qualitätsvolle Veränderung des Bildungsangebotes sowie die zunehmende Professionalisierung der Lehrkräfte zum Ziel hat. Dabei lassen sich Schulentwicklungsprozesse sowohl auf der Organisations-, als auch auf der Personal- sowie auf der Unterrichtsebene beschreiben. Zwei Grundannahmen gehen diesem Verständnis von Schulentwicklungstheorien voraus: Zum einen wird Schule als Lern- und Lebensraum verstanden, als ein „Haus des Lernens“ (Bildungskommission NRW 1995), in dem sich alle Personen als Lernende verstehen. Damit geht zum anderen ein Verständnis von Schulen als „Lernende Organisationen“ einher, die aufgrund ihrer systemischen Eigenschaften in der Lage sind mit hoher Innovationsbereitschaft auf sich wandelnde gesellschaftliche Herausforderungen angemessen reagieren zu können, und zwar dergestalt, dass notwendige Bildungsprozesse bei den Schülerinnen und Schülern angestoßen werden (Rahm 2005, S.25 sowie S. 45). Das Anstoßen dieser Bildungsprozesse erfordert ein didaktisches Umdenken der Lehrkräfte, was neben der Umgestaltung von Lehr- und Lernarrangements auch die Veränderung der pädagogischen Arbeitsweise miteinander nach sich zieht (Risse 2003). Weiterhin finden diese in das Gesamtsystem Schule eingebetteten Prozesse der Personal- und Unterrichtsebene zusätzlich in Interaktion mit der umgebenden Umwelt und den darin vorhandenen Ressourcen statt.

*„Schulentwicklung wird demnach konsequent zu einem Lernprozess, innerhalb dessen die Institution sich in interaktiver Vernetzung mit ihrer Umwelt verändert, sprich: Qualität entwickelt.“
(Rahm 2005, S. 25)*

Der angesprochene Lernprozess kann somit alle beteiligten Personen erfassen und organisationales Lernen im Sinne einer Entwicklung der Einzelschule ermöglichen.

2.3 Relevanz von Organisations- und systemtheoretischen Perspektiven für Schulentwicklung

Mit dem Einzug organisationstheoretischer Perspektiven in die Schul- und Unterrichtsforschung haben sich Ansätze zum Verständnis von Schule als „Lernende Organisation“ (Senge 1998; Argyris & Schön 1999) sowie vielfältige Diskurse zur Organisationsentwicklung oder zum Organisationslernen in schu-

lischen bzw. pädagogischen Einrichtungen (Eichler 2008) etabliert. Gemeinsam ist diesen Ansätzen, dass Organisationen wie Schulen die Fähigkeit zugesprochen wird, Entwicklungspotenzial zu entdecken und Veränderungen eigentätig voran zu treiben.

Organisationen werden von Rosenstiel (2007) als von Menschen geschaffene Systeme beschrieben, die Bedeutung für ihre Mitglieder durch ihre Wahrnehmung, Deutung und Interpretationen gewinnen. Sie setzen sich aus Individuum und Gruppe zusammen und weisen bestimmte Strukturen zur Koordination einzelner Tätigkeiten auf. Im wissenschaftlichen Diskurs existieren verschiedene Organisationsbegriffe (Schreyögg 2008). Unterschieden wird mehrheitlich zwischen dem instrumentellen und dem institutionellen Organisationsbegriff. Der instrumentelle Organisationsbegriff – ein Unternehmen *hat* eine Organisation – meint die existierenden Abläufe, Hierarchien und Strukturen. Im Gegensatz dazu wird mit dem institutionellen Organisationsbegriff – ein Unternehmen *ist* eine Organisation – der Blick auf das gesamte System, auf die Institution gelenkt und demnach die Organisation als ein zielgerichtetes, offenes soziales Gebilde mit einer formalen Struktur bezeichnet. Eine Organisation ist folglich ein System interindividueller Beziehungen zwischen Individuen und Gruppen mit gemeinsamer Ziel- und Zwecksetzung, geregelter Arbeitsteilung und beständigen Grenzen. Indem der institutionelle in Abgrenzung zum instrumentellen Organisationsbegriff eben nicht nur die Abläufe, Hierarchien und Strukturen, sondern das gesamte soziale System zu erfassen sucht, greift dieser Begriff weiter. Diese weite Perspektive mit den markierten Elementen des institutionellen Organisationsbegriffs (ebd.) scheint im Hinblick auf den Untersuchungsgegenstand schulischer Systeme angemessen.

Versteht man Schule nun als eine Organisation mit gemeinsamer Ziel- und Zwecksetzung, geregelter Arbeitsteilung und beständigen Grenzen, so finden sich die genannten Merkmale auf verschiedenen Ebenen wieder: Schulen sind zunächst vom Menschen geschaffene und gestaltete Organisationen mit der historisch gewachsenen Ziel- und Zwecksetzung der Einhaltung des gesellschaftlichen Bildungsauftrages sowie mit einer bestimmten, in sich geschlossenen formalen Struktur (dies ließe sich bspw. in Organigrammen zur Personalstruktur, aber auch in ihren zeitlichen und arbeitsteiligen Abläufen abbilden). Weiterhin sind Schulen sowohl in sich geschlossene Systeme mit wiederum in sich geschlossenen Subsystemen als auch ihrer Umwelt gegenüber offene Systeme. Nicht nur im Zusammenhang mit der Diskussion um die In-

tegration externen Lehrpersonals als Kooperationspartner⁵ sollte der Diskurs um eine Öffnung des schulischen Systems bedeutsam werden.

Der Einfluss der organisationstheoretischen sowie systemischen Sichtweisen im Forschungsprozess sowie in der Ergebnisdarstellung zeigt sich wie folgt: Zunächst ist das Gesamtsystem Schule mit seinen formalen Regeln, Strukturen und gesetzten Bedingungen in den Blick genommen worden. Zusätzlich dazu sind die sozialen Prozesse zwischen Individuen und Gruppen innerhalb und außerhalb der Lehr-Lernsettings unter Analyseaspekten wie Kooperationsstrukturen, Klimaentwicklungen, Lern- und Entwicklungsprozessen bei Schülerinnen und Schülern und Lehrpersonal betrachtet worden. Diese Gesichtspunkte haben sich in der Erstellung und Anwendung des schulischen Erklärungsmodells niedergeschlagen (vgl. *Abschnitt 4*).

2.3.1 Organisationskultur

Im organisationstheoretischen Diskurs ist die Erforschung und Analyse einer spezifischen *Organisationskultur* prominent. Ihr wird derzeit vor allem im Hinblick auf Innovation und Veränderung von Organisationen große Bedeutung beigemessen. Hier haben sich konzeptionelle Ansätze etabliert, die sich auf verschiedene Weise dem Konstrukt der Organisationskultur annähern. Für Schul- und Unterrichtsforschung, die sich der Untersuchung von Schulen in der Gründungs- bzw. frühen Entwicklungsphase widmet, erscheint ein Einbezug dieser Perspektiven unerlässlich.

„Drei-Ebenen-Modell“ der Organisationskultur

Als einer der wichtigsten Vertreter in der Erforschung und Beschreibung der Organisationskultur gilt Edgar H. Schein (1985; dt. 1995). Er hat ein „Drei-Ebenen-Modell“ der Organisationskultur aufgestellt und beschreibt darin das Zusammenspiel dreier Kulturschichten. Auf der obersten Schicht wird die Or-

5 Im Rahmen der StEG-Untersuchung werden derzeit unter der Bezeichnung „Kooperationspartner“ die übergreifenden kooperierenden Institutionen o. ä. gefasst, die konkreten Lehrenden jedoch als „weiteres pädagogisch tätiges Personal“ bezeichnet. Die Autorin ist sich dieser Differenzierung bewusst, möchte aber aus Gründen der besseren Lesbarkeit und Verständlichkeit im vorliegenden Beitrag die außerschulischen Lehrenden stets als Kooperationspartner bezeichnen. Des Weiteren wird aus gleichen Gründen auf die gesonderte Ausweisung der weiblichen Form verzichtet, es werden jedoch stets beide Geschlechter berücksichtigt.

ganisationskultur in Artefakten, Objekten oder Verhaltensweisen sichtbar. Dazu gehören bspw. bestimmte Dokumente und die räumliche Ausgestaltung, aber auch Rituale, Kleidung oder Slogans. Auf der mittleren Schicht finden sich die kollektiven Werte der Organisationsmitglieder, die das Verhalten der Organisationsmitglieder wesentlich steuern. Die dritte und gleichzeitig tiefste Ebene umfasst diejenigen Grundannahmen der Organisationsmitglieder, die sich im Unterbewusstsein der Organisation befinden und als langfristig konstante Auffassungen gelten. Sie werden nicht diskutiert und haben ebenfalls großen Einfluss auf das Verhalten der Organisationsmitglieder. Hierzu zählen bspw. die Beziehung zur Umwelt, das Wesen der sozialen Beziehungen untereinander oder auch das in der Organisation vorherrschende Menschenbild.

Organisationskultur in der Organisations- und Unternehmenskulturforschung

Im Diskurs um Funktion und Entwicklung der Organisationskultur nimmt u. a. die *Organisations- und Unternehmenskulturforschung* eine bedeutsame Position ein. Vertreter⁶ dieser Forschungsperspektive gehen prinzipiell von folgenden drei Annahmen aus: Organisationskultur sei erstens nur eine von mehreren organisationsinternen Variablen. Zweitens bestehe die Organisationskultur aus Teilelementen, die sich zu einem homogenen Ganzen zusammenfügen. Schließlich erfülle die Organisationskultur in dieser Auffassung mehrere bestandswichtige Funktionen, die für die Zielerreichung, die Innovationsfähigkeit und den Markterfolg eines Unternehmens von zentraler Bedeutung sind. Die Organisationskultur bestehe demgemäß aus einer Vielzahl von verschiedenen Produkten und Artefakten, aber auch aus Mythen oder Sagen, aus Symbolen, Werten und Normen. Organisationskultur wird in diesem Verständnis als System der Gewährleistung und Aufrechterhaltung „gemeinsamer Werte und Normen“ angesehen. Weiterhin nimmt man an, dass eine starke Organisationskultur positive Auswirkungen auf organisatorische Variablen wie Effizienz oder Leistung habe. Dementsprechend müsse das Ziel sowohl die Stärkung der Organisationskultur als auch die Offenlegung und Messung angenommener Ursache-Wirkungs-Zusammenhänge sein (Franzpötter 1997, S. 24ff.).

6 Zitiert nach Franzpötter, der sich dabei auf Autoren wie Peters & Waterman (1983), Heinen (1987) Scholz & Hofbauer (1990) beruft.

Organisationskultur als „Interpretations- und Verständigungsrahmen“

Franzpötter (1997) nähert sich der Organisationskultur aus sozialwissenschaftlicher Perspektive. Er widerspricht der Annahme, Organisationskultur sei ausschließlich als ein gemeinsames Wertesystem zu verstehen. Vielmehr wird Organisationskultur als „Geflecht von Bedeutungen, in denen Menschen ihre Erfahrungen interpretieren und nach denen sie ihr Handeln ausrichten“ betrachtet (Geertz 1983, S.99, zitiert nach Franzpötter 1997, S. 36). Weiterhin beschreibt Franzpötter die Organisationskultur als:

„[...] dasjenige, [...] was in einem konkreten Organisationssetting als signifikant und relevant gilt, was als positiv oder negativ bewertet, als Denk- und Interpretationsweise der eigenen Geschichte, der Organisationsmitglieder und der Organisationsumwelt institutionalisiert ist.“ (Franzpötter 1997, S. 59)

Außerdem versteht er Organisationskultur als sozial anerkannte und symbolisch vermittelte *Wirklichkeitsinterpretation*, die in Wechselwirkung mit der Umwelt der Organisation in der täglichen Handlungspraxis eines Organisationssettings entsteht. Sie verleiht dem Handlungsgeschehen Struktur, und zwar durch institutionalisierte Deutungsmuster, Überzeugungen, Grundannahmen, Gefühle oder auch Selbstverständlichkeiten. Im Gegensatz zur Annahme von Organisationskultur als gemeinsames Werte- und Normensystem versteht Franzpötter Organisationskulturen als *Interpretations- und Verständigungsrahmen*, die nicht einfach existieren oder aufgestellt werden können. Diese Orientierungsrahmen müssen permanent von den Akteuren eines Organisationssettings in gemeinsamen „Handlungs-, Erfahrungs- und Interpretationsprozessen geschaffen, institutionalisiert und legitimiert werden“ (Franzpötter 1997, S. 59).

Zusammenfassung

Organisationskultur wird stets als etwas Zusammengesetztes verstanden, bestehend aus verschiedenen Teilelementen wie Artefakten, Werten, Normen oder Grundannahmen. Diese sind zur Bestimmung und Ausbildung einer spezifischen Organisationskultur wichtig und werden jeweils von den Organisationsmitgliedern geteilt. Während die funktionalistische Sichtweise die Organisationskultur als *Variable* betrachtet mit der Funktion der besseren Zielerreichung einer Organisation, der Verbesserung ihrer Innovationsfähigkeit oder ihres Markterfolges, betonen die systemische und die sozialwissenschaftliche Perspektive, dass es sich bei der Organisationskultur um eine *Eigenschaft*

(Klimecki & Probst 1990, zitiert nach Metz-Göckel 1996) oder um einen *Verständigungs- und Interpretationsrahmen* (Franzpötter 1997) handelt. Dieser umfassende Rahmen dient dazu, Erfahrungen angemessen zu interpretieren und Handlungsweisen entsprechend zu generieren. Nach Franzpötter ist dieser Orientierungsrahmen nicht automatisch gesetzt, sondern muss stets neu ausgehandelt werden und zwar aufgrund gemeinsam geteilter Handlungs-, Erfahrungs- und Interpretationsprozesse durch die Mitglieder. Außerdem entsteht diese Wirklichkeitsinterpretation durch Wechselwirkungen zwischen Organisation und Umfeld. Die oben angesprochenen Normen und Deutungsmuster verleihen diesem spezifischen Handlungsgeschehen Struktur.

Im Hinblick auf schulische Veränderungsprozesse steht nicht vorrangig und ausschließlich die Effektivitäts- oder Leistungssteigerung eines schulischen Systems, basierend auf der Offenlegung einfacher Ursache-Wirkungs-Zusammenhänge, im Vordergrund. Neben der Aufrechterhaltung eines gemeinsam geteilten Norm- und Wertesystems werden in schulischen Systemen v. a. Wechselwirkungen einerseits innerhalb der Organisation zwischen Individuen oder zwischen verschiedenen Gruppen und andererseits zwischen Organisation und Umwelt bedeutsam, die zur Ausbildung des spezifischen Orientierungsrahmens führen. Dieser Rahmen ermöglicht, dass die schulischen Akteure ihre individuellen Erfahrungen angemessen interpretieren und ihre Handlungsweisen entsprechend daran ausrichten können. Gemeinsame Handlungs-, Erfahrungs- und Interpretationsprozesse bilden dafür die Grundlage und sind natürlicher und notwendiger Bestandteil von Schule und Unterricht.

2.3.2 Die Bedeutung von Schul- und Lernkultur für Schulentwicklungsprozesse

Schul- und Lernkultur

Zum einen steht in Anlehnung an organisationstheoretische Perspektiven die Offenlegung einer schulspezifischen Organisationskultur – oder anders: Schulkultur – im Mittelpunkt. Schulkultur gibt in diesem Verständnis den übergreifenden *Interpretations- und Verständigungsrahmen* (Franzpötter 1997) vor.

Im Rahmen der von der Bertelsmannstiftung geförderten Initiative „SEIS – Selbstevaluation in Schulen“ (*siehe auch*: Niedersächsisches Landesamt für Lehrerbildung, n. d.) gehört *Schulkultur* (im Zusammenhang mit Schulklima) neben den Bereichen *Lernen und Leben*, *Bildungs- und Erziehungsauftrag*,

Führung und Management sowie *Zufriedenheit* zu den fünf generierten Qualitätsbereichen von Schulentwicklung.

„Es [Anm.: in die Definition zur Schulkultur] ist vielmehr das gesamte Persönlichkeitsspektrum der Agierenden einzubeziehen. Die Schulkultur ist dann das gelebte Selbstverständnis der beteiligten Menschen, das einerseits von deren Aktionen geschaffen wird, das andererseits bildend (!) auf die Persönlichkeiten (und die Strukturen) zurückwirkt.“ (Fuchs 2009, S. 5)

Schulkultur kann nicht einfach von außen gesetzt werden, sondern muss durch die Interaktion der beteiligten Personen immer wieder neu ausgehandelt werden. Gemäß dieser Auffassung wirkt der Orientierungsrahmen auf die am Schulleben beteiligten Akteure sowie auf die strukturellen Bedingungen zurück.

Zum anderen stellt der Unterricht in den verschiedenen Lehr-Lernsituationen das Kernelement von Schule dar und sollte dementsprechend im Hinblick auf die Entwicklung einer veränderten *Lernkultur* mit besonderer Sorgfalt bedacht werden. Risse (2003) formuliert u. a. wie folgt:

„Das Ziel der Schule, Kinder und Jugendliche auf ihrem Weg zur Ausformung einer eigenen Persönlichkeit fördernd und fordernd zu begleiten und sie zu befähigen, am Leben in der Gesellschaft teilzunehmen, kann nur erreicht werden, wenn schulische Lernprozesse in einer „Lernkultur“ aufgehoben sind, die das Erreichen dieser Ziele unterstützt.“ (Risse 2003, S. 5)

Risse thematisiert wichtige Grundannahmen einer Lernkultur aus systemischer Perspektive, die für das vorliegende Forschungsvorhaben diskussionsleitend sind. Sie generiert vier Leitprinzipien einer systemischen Schulentwicklung. An erster Stelle steht (1) die Verbindlichkeit eines pädagogischen Leitbildes, wie es bspw. in Schulprogrammen verankert wird. Weiterhin benennt sie (2) die „systematische Erarbeitung von fachbezogenen und fachübergreifenden Lernarrangements, die in immer stärkerem Maße selbstregulierte Lernprozesse ermöglichen“. Außerdem sollte (3) eine Verständigung darüber erfolgen, wie der „Ausgleich zwischen gesetzten Standards und individuellen Leistungsdispositionen der einzelnen Schülerinnen und Schüler erreicht werden kann“. Als viertes Leitprinzip schließlich formuliert Risse (4) die Zusammenführung von pädagogischem und organisatorischem Management, was vor allem in der Verantwortung der Schulleitung liegt. Diese vier Prinzipien seien

nicht losgelöst voneinander zu betrachten, sondern überschneiden sich vielfach.

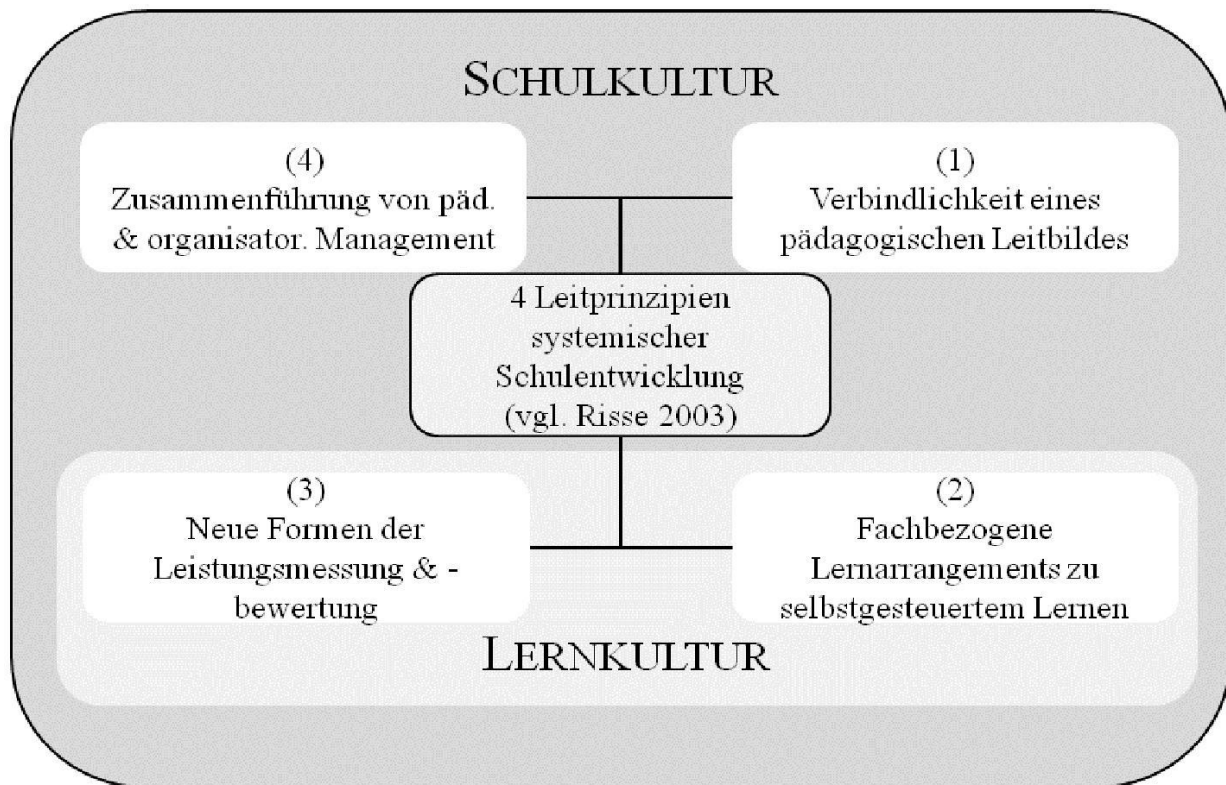


Abb. 1: Darstellung der vier Leitprinzipien systemischer Schulentwicklung nach Risse
Quelle: eigene Darstellung nach Risse (2003)

Im Sinne der o. g. Darlegungen werden in der vorliegenden Betrachtung die Prinzipien eins und vier als Elemente der Entwicklung einer Schulkultur (i. S. eines Orientierungsrahmens) verstanden, während die Prinzipien zwei und drei den Unterricht und die damit verbundenen notwendigen neuen Lehr- und Lernformen fokussieren. Dementsprechend sollen sie als Elemente einer Lernkultur diskutiert werden. Allerdings sind sie, wie in *Abbildung 1* ersichtlich, in die umfassende Schulkultur eingebettet. Für die vorliegende Arbeit soll zwischen einer Schul- und einer Lernkultur unterschieden werden.

3 Methodologie des Forschungsprozesses

Die methodologische Ausrichtung der vorliegenden Arbeit orientiert sich an folgenden Schwerpunkten: Erstens entspricht das Gesamtforschungsdesign einer *qualitativen Fallstudie*⁷. Zweitens sind für ein umfassendes Verständnis

⁷ Dabei wird einem Verständnis von Fallstudien im qualitativen Paradigma von

Grundannahmen und Verfahren der *qualitativen Evaluation* angewandt worden. Schließlich ist für die Analyse und Auswertung des Datenmaterials dem methodischen Vorgehen des sogenannten *Thematischen Kodierens* nach Uwe Flick (2007) in enger Verzahnung zu Verfahrensschritten der *Grounded Theory Methodology*⁸ gefolgt worden. Einen weiteren Schwerpunkt bildet die *Triangulation*, die als übergreifendes methodisches Prinzip im Sinne von Perspektiven- und Erkenntniserweiterung eingesetzt worden ist (Flick 2008). Dementsprechend sind folgende Einzelmethoden der Erhebung und Analyse von Daten ausgewählt und aufeinander abgestimmt worden: Die Datenerhebung erfolgte mittels leitfadengestützter Interviews in Zusammenhang mit der nicht-teilnehmenden, aber direkten Beobachtung ausgewählter Situationen sowie der ergänzenden Dokumentenanalyse. Während mit der Schulleitung, den Lehrenden und den Eltern sogenannte theoriegenerierende Experteninterviews geführt worden sind (Bogner & Menz 2005), sind im Rahmen der Schülerbefragung fokussierte Interviews (Reinders 2005) zum Einsatz gekommen. Das bereits genannte Thematische Kodieren schien zur fallspezifischen sowie fallübergreifenden Analyse der gewonnenen Daten geeignet (Naacke & Lehmann-Wermser 2009; Naacke, *in Druck*).⁶

4 Ergebniskomplex I: Das schulische Erklärungsmodell

Im Forschungsprozess aber auch in der Darstellung seiner Ergebnisse sind zunächst das Gesamtsystem Schule mit seinen formalen Regeln, Strukturen und gesetzten Bedingungen in den Blick genommen worden. Weiterhin sind die sozialen Prozesse zwischen Individuen und Gruppen unter Analyseaspekten wie Kooperationsstrukturen, Klimaentwicklungen, Lern- und Entwicklungsprozessen bei Schülerinnen und Schülern und Lehrpersonal betrachtet worden. Demzufolge sind mit der Erstellung eines schulischen Erklärungsmodells mehrere Überlegungen verbunden. Einerseits ist es ein forschungsmethodologischer Schritt. Denn mit Schaubildern oder Modellen lässt sich die Komplexität der o. g. Aspekte übersichtlich darstellen. Andererseits kann auf solcherart

Autorinnen und Autoren wie John W. Creswell (2005), Siegfried Lamnek (2005), Sharan B. Merriam (1988), Robert E. Stake (2000) und Robert K. Yin (2005) gefolgt.

- 8 Für weiterführende Informationen zur Grounded Theory Methodology siehe u. a. Strauss & Corbin 1996 sowie Strauss 1998.

Modelle als weiterführende Erklärungsinstrumente auch in anderen schulischen Kontexten zurück gegriffen werden.

Sowohl im Rahmen von qualitativen Fallstudien als auch im Rahmen der Grounded Theory Methodology wird auf die Verwendung von Modellen und Grafiken im Analyse- und Auswertungsprozess hingewiesen (Strauss & Corbin 1996; Strauss 1998; Yin 2005).

“The use of logic models as an analytic technique consists of matching empirically observed events to theoretically predicted events.” (Yin 2005, S. 127)

Yin differenziert diesbezüglich vier solcher *Logischer Modelle*, die auf verschiedenen Ebenen die Ergebnisdarstellung strukturieren können. Neben Modellen, die auf Individualebene angesiedelt sind, beschreibt er Modelle, die Beschreibungen auf Organisationsebene zulassen. Hier unterscheidet er zwischen linearen und dynamisch zirkulierenden Modellen. Schließlich erwähnt Yin Modelle, die zur Veranschaulichung von bestimmten Programmen hinzugezogen werden können. Im vorliegenden Fall sind es die von Yin benannten *Logischen Modelle auf Organisationsebene* und zwar im Besonderen die dynamischen Modelle. Denn sie sind geeignet, Veränderungsprozesse in Organisationen, z. B. von Schulen zu verdeutlichen (Yin 2005, S. 130).

Strauss (1998) schlägt die Anwendung von Diagrammen vor. Sie können in allen drei Phasen des Kodierens eingesetzt werden und sind durch Memos und andere theoretische Überlegungen anzureichern. Zudem unterscheidet Strauss zwischen logischen und integrativen Diagrammen. Während logische Diagramme „die Entfaltung der logischen Beziehungen zwischen Kategorien und Subkategorien“ aufzeigen (Strauss & Corbin 1996; S. 167), werden integrative Diagramme als „visuelle Darstellungen analytischer Gedankengänge zum Ausprobieren und Demonstrieren von konzeptuellen Verknüpfungen“ angesehen (ebd.). Erstere beschreiben die logische Verknüpfung der Kategorien untereinander und letztere sind offen für eigene Vorstellungen und Einfälle. Sie sind zudem ein wichtiges Hilfsmittel auf dem Weg zu einer integrierten Theorie.

„Das integrative Diagramm hilft dem Forscher, sich ein deutlicheres Bild davon zu machen, wie weit er in seiner Forschungsarbeit nach Datenerhebung, Kodieren und Memoschreiben gekommen ist. [...] Von Zeit zu Zeit sollte man sich das aktuelle Diagramm anschauen und sich die Frage stellen: Was erfaßt dieses

Diagramm noch nicht, was ich jetzt weiß? Wenn sich dann eine bestimmte Wissensmenge angesammelt hat, ist es an der Zeit, ein neues Diagramm zu entwerfen.“ (Strauss 1998, S. 239)

Im Folgenden soll das Erklärungsmodell allgemein vorgestellt werden. Es hat sich im Zuge der induktiven Analyseverfahren, ausgehend von MUKUS entwickelt. Das Modell bildet als Rekonstruktion der Sicht- und Handlungsweisen der befragten Personen sowie der übrigen erhobenen Daten den gewählten Wirklichkeitsausschnitt ab. Durch Prozesse der Re-Kodierung⁹ sowie durch die Beschäftigung mit Fachliteratur und der intensiven Auseinandersetzung mit der hier konkreten Einzelfallschule hat sich das Modell zwangsläufig verändert. Dabei folgten die beschriebenen Phasen nicht einem linearen Nacheinander, sondern verstehen sich als ein Hin- und Herpendeln zwischen Modell und weiterem Material mit dem Ziel einer reichhaltigen Abbildung und strukturierten Erläuterung der erhobenen Daten. Bereits im Forschungsverlauf von MUKUS sind datenbezogen über die teilnehmenden Einzelschulen hinweg *Große Themen* generiert worden, die als Hauptkategorien im Sinne der Grounded Theory Methodology verstanden werden. Diese Großen Themen fügen sich in das schulische Erklärungsmodell ein. Indem die Hauptkategorien an den Einzelschulen unterschiedliche Bedeutung einnehmen sowie verschiedentlich beeinflusst werden können, bleibt das Modell individuell dynamisch und veränderbar. Für die theoretische und modellbezogene Weiterentwicklung sollen zwei Modelle eingefügt werden: *Modell A* illustriert den Erkenntnisstand der MUKUS-Studie, wohingegen sich *Modell B* auf die Veränderungen im Zusammenhang mit der vorliegenden Arbeit stützt.

9 Während der Begriff der Re-Kodierung vornehmlich im Rahmen quantitativ-statistischer Verfahren gebraucht wird, und diesbezüglich die Umformulierung der erhobenen Antworten in Zahlenformate meint, die die Grundlage für die Berechnungen bilden, ist der Begriff hier im Sinne einer Rücküberprüfung der Kategorien in den Daten zu verstehen, wie er von Strübing (2008) verwendet wurde. Dieses Verfahren der Rücküberprüfung hat die theoretische Schließung zum Ziel.

Modell A

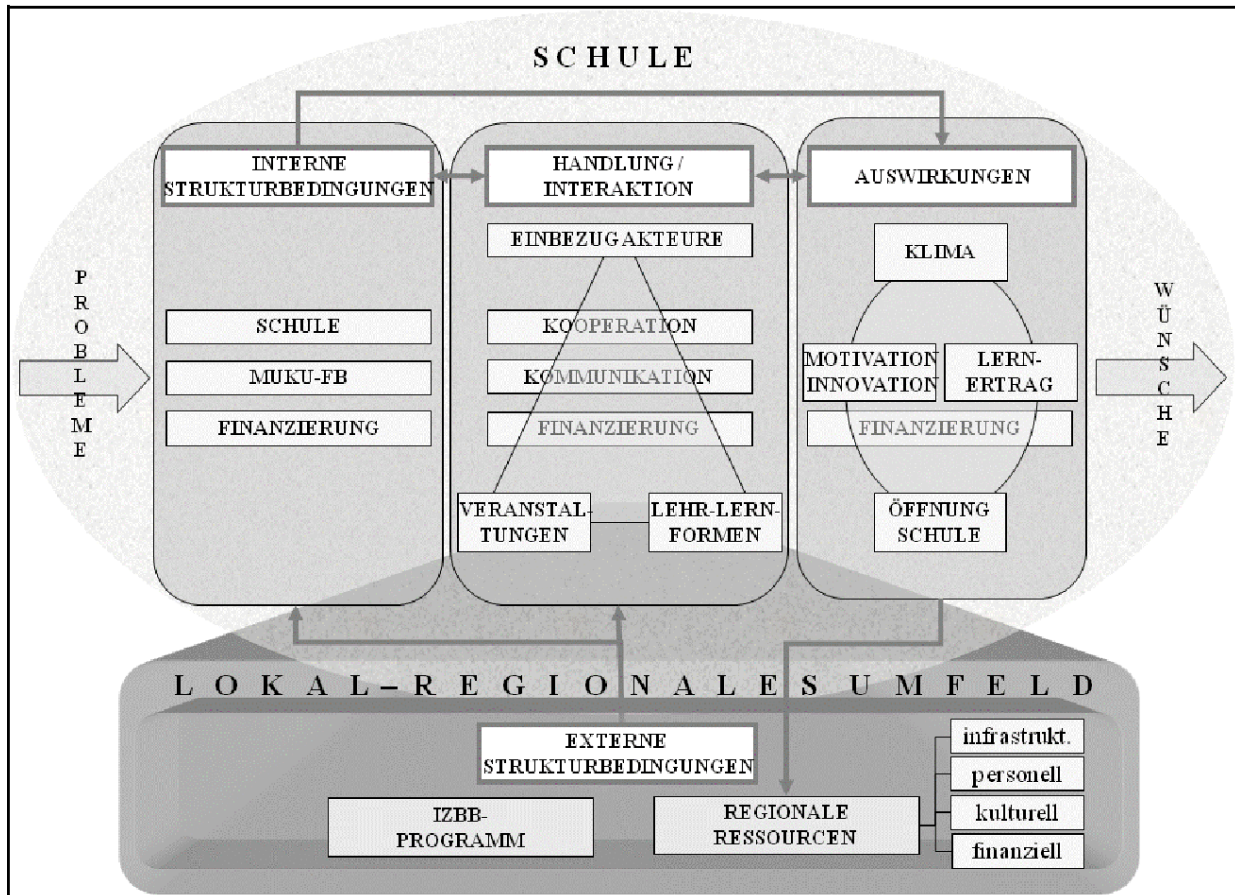


Abb. 2: Allgemeines schulisches Erklärungsmodell

Anmerkungen:

Verwendete Kürzel: MUKU-FB = musisch-kulturelle Fachbereich; IZBB-Programm = „Investitionsprogramm Zukunft, Bildung und Betreuung“

Die Pfeile in eindimensionaler Richtung drücken aus, dass A von B beeinflusst wird (bzw. B auf A einwirkt). Die Pfeile mit zweifacher Richtungsausprägung verdeutlichen, dass A und B in gegenseitiger Wechselwirkung stehen.

Quelle: eigene Darstellung MUKUS 2009 (Susanne Naacke)

Modell B

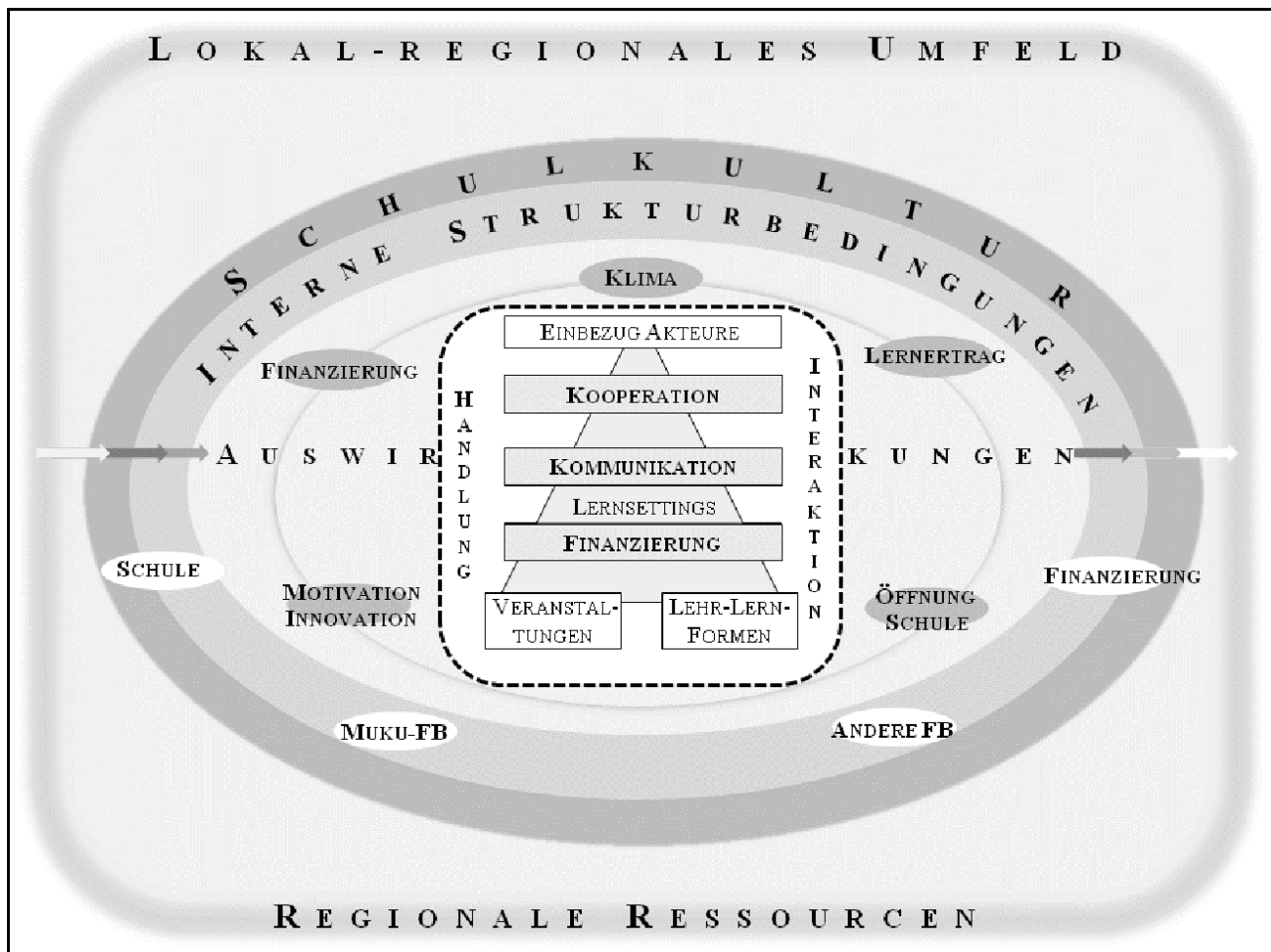


Abb. 3: Allgemeines Schulmodell

Anmerkungen: Verwendete Kürzel: MUKU-FB = musisch-kultureller Fachbereich, ANDERE FB = andere Fachbereiche
 Quelle: eigene Darstellung 2010 (Susanne Naacke)

Wie in *Abbildung 3* zum Modell B dargestellt, symbolisiert die äußere Ellipsenform die Schule. Sie ist in ein konkretes lokal-regionales Umfeld eingebettet, welches durch bestimmte Voraussetzungen oder externe Strukturbedingungen gekennzeichnet ist. Als Voraussetzungen, die auf die Gestaltung des schulischen Lebens Einfluss nehmen können, werden regionale Ressourcen verstanden, und zwar finanzieller, infrastruktureller, personeller, kultureller sowie kirchlicher Art. Weiterhin verdeutlicht das Modell, dass das System Schule von einer spezifischen Schulkultur geprägt ist. Die internen Strukturbedingungen umfassen die Einflussgrößen Schule, musisch-kultureller Fach-

bereich¹⁰, Finanzierung sowie die anderen Fachbereiche. Als zentrales Verbindungsglied erweist sich der Bereich der Handlungs- und Interaktionsstrategien mit den Hauptkategorien fachspezifische Lehr- und Lernformen, zusätzliche Veranstaltungen o. a. Präsentationsformen, Einbezug der (schulischen) Akteure, sowie Kooperation, Kommunikation und erneut Finanzierungsaspekte. Im Bereich der Auswirkungen werden neben finanziellen Entwicklungen v. a. Veränderungen bzgl. des Schul- und Unterrichtsklimas, des Lernertrags auf Schülerseite, bzgl. Motivation / Innovation auf Seiten des lehrenden Personals sowie bzgl. einer Öffnung der Schule thematisiert.

4.1 Vergleich von Modell A und B

Beide Modelle fungieren als Abbildung des Systems Schule mit der besonderen Verankerung des musisch-kulturellen Fachbereichs. Sie geben Auskunft über Struktur- und Kontextbedingungen, Handlungs- und Interaktionsstrategien sowie Aus- und Rückwirkungen einerseits innerhalb des Systems Schule und seiner Subsysteme und andererseits zwischen Schule und den äußeren Einflussfaktoren. Während beim Modell A der MUKUS-Studie die verschiedenen Struktur- und Handlungsbereiche noch relativ geschlossen nebeneinander standen, schien es nun notwendig, die Interaktionen und die Wechselwirkungen mittels durchlässiger Ellipsen stärker zu verdeutlichen. Außerdem ist die Kernkategorie - das zentrale Interaktionsdreieck - in den Mittelpunkt gestellt worden, um damit die wechselseitige Steuerung von Schul- und Lernkultur deutlicher zum Ausdruck zu bringen. Insgesamt weisen beide Modelle zwar die gleichen Hauptbereiche mit den jeweiligen Großen Themen auf, allerdings erfahren sie eine veränderte Anordnung im Modell B. Zudem konnte im Laufe der Analysearbeit ein weiteres Großes Thema - die Schulkultur als Interpretations- und Verständigungsrahmen - generiert werden. Auch bei den internen Strukturbedingungen treten modellbezogene Abweichungen auf: Im Modell B ist zusätzlich zu den Einflussgrößen *Schule*, *musisch-kultureller Fachbereich* und *Finanzierung* die Hauptkategorie *andere Fachbereiche* auf-

10 Dass modellbezogen von musisch-kulturellen Praxen gesprochen wird, hat seine Ursache in der MUKUS-Studie, die alle drei künstlerischen Fächer erforscht hat. Die vorliegende Arbeit ist in die umfassenden Ergebnisse von MUKUS eingebettet, sodass in der allgemeinen Ergebnispräsentation zum Modell stets auf den gesamten musisch-kulturellen Bereich Bezug genommen wird. Für die konkrete Auseinandersetzung mit dem evangelischen Domgymnasium gelten die Erkenntnisse vorwiegend für das Fach Musik.

genommen worden. Dieser Befund erklärt sich daher, dass bei den MUKUS-Schulen ausschließlich der musisch-kulturelle Bereich fokussiert betrachtet worden ist, während beim Dissertationsprojekt eine stärker ganzheitliche Betrachtung vollzogen wurde. Weiterhin erweisen sich bei beiden Modellen die Handlungs- und Interaktionsstrategien als zentrales Verbindungsglied: Das *Interaktionsdreieck* (vgl. nachfolgenden Abschnitt) bildet jeweils das Zentrum des Modells und ist charakteristisch für musisch-kulturelle und insbesondere musikalische Praxen. Denn diese erfordern und ermöglichen im Sinne einer erfolgreichen Umsetzung den Dreischritt aus vielfältigen Lehr- und Lernformen, zusätzlichen Veranstaltungen und dem Einbezug der jeweiligen Akteure. Schließlich werden gleichermaßen Auswirkungen auf Individual-, Gruppen- und Schulebene thematisiert.

Die Kernkategorie als Erklärungsschlüssel

Im Sinne der Grounded Theory Methodology wird sich eine Kategorie als zentral erweisen, die sämtliche Informationen unter ihrem Erklärungszusammenhang bündelt. Im vorliegenden Fall ist es nicht eine einzelne Kategorie, sondern es ist das spezifische Zusammenkommen von besonderen Lehr- und Lernformen, zusätzlichen Präsentationen und dem Einbezug der Akteure. Dieses Zusammenspiel wird Interaktionsdreieck genannt (vgl. *Abbildung 4*).

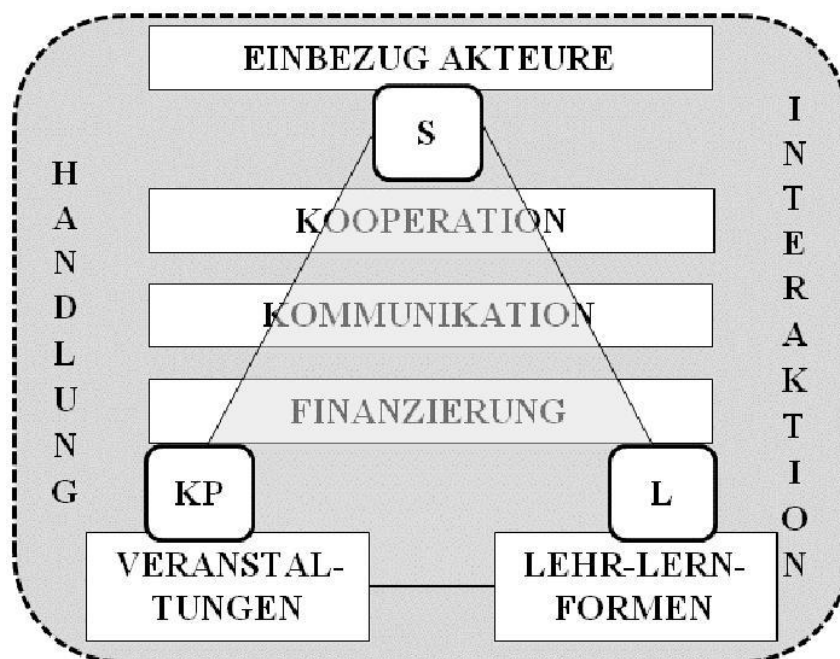


Abb. 4: Kernkategorie — das Interaktionsdreieck

Quelle: eigene Darstellung 2010 (Susanne Naacke)

Es ist nicht nur notwendig für musikalische Praxen, sondern musikalische Praxen bringen dieses Dreieck typischerweise auch hervor. Erst das Aufeinandertreffen der drei Kategorien und ihr Zusammenwirken ermöglichen vielfältige positive Befunde auf Individual-, Gruppen- und Schulebene. Nicht nur bei MUKUS, auch beim Domgymnasium haben die befragten Akteure immer wieder dieses Dreierverhältnis als Erklärungsmuster herangezogen. Es lässt sich zudem formulieren, dass jener beschriebene Wechsel aus Klein- und Großgruppenunterricht oder anderen Lehr-Lernformen ergänzt durch die zusätzlichen Veranstaltungen (wie Schulkonzerte, Aufführungen im lokalen Umfeld etc.), die regelmäßig im Schuljahresverlauf am Ende einer bestimmten Phase stehen, in besonderer Weise charakteristisch ist für musikalische Praxen. Erst durch diesen regelmäßigen Wechsel wird ein intensiver Einbezug der Akteure ermöglicht. Zudem ist der Einbezug der Akteure unerlässlich, weil musisch-kulturelle Praxen nur deshalb realisierbar sind, weil sich jeder mit seinen Fähigkeiten einbringt.

Einbezug der Akteure

Eigenengagement und Partizipation der beteiligten Akteure hinsichtlich musisch-kultureller Praxen sind sowohl notwendige Voraussetzung als auch zentrale Auswirkung und werden unter der Hauptkategorie „Einbezug Akteure“ zusammengefasst.

Unterschiedliche Konzepte oder Situationen verlangen unterschiedliche Einbindung der Akteure. Die Frage der Einbindung von Lehrkräften, Kooperationspartnern sowie Schülerinnen und Schülern im schulischen Kontext ist logisch und einleuchtend. Dagegen erscheint die Einbindung von Eltern und Schulleitung zunächst etwas befremdlich. Allerdings ist deren Einbezug von immenser Bedeutung. Musisch-kulturelle Praxen sind häufig mit einem hohen Zeit- und Arbeitsaufwand sowie Kosten verbunden. Fehlten Engagement und Unterstützung durch das Elternhaus oder die generelle Akzeptanz für den oft hörbaren Arbeitsaufwand (bspw. Üben von Instrumenten oder Theaterrollen), wäre die Durchführung v. a. für die beteiligten Schülerinnen und Schüler gefährdet. Ähnlich verhält es sich mit dem Einbezug der Schulleitung. Musisch-kulturelle Praxen sind sowohl im Schulleben als auch zur Darstellung der Schule im städtischen Umfeld gefragt: bei Ausgestaltungen schulischer (oder städtischer) Veranstaltungen dürfen Chor, Schulband oder auch die Theater-AG – so vorhanden – nicht fehlen. Schulleitungen, die die angesprochenen musisch-kulturellen Konzepte und Inhalte mit ihrem Engagement unterstützen, schaffen bspw. für die meisten Akteure günstig gelegene Zeitfenster für die

Teilnahme an Arbeitsgemeinschaften oder Ganztagsangeboten, treten in Kooperationsverhandlungen mit außerschulischen kulturellen oder anderen Institutionen oder sie schaffen die notwendigen räumlichen und finanziellen Grundlagen. Zusammengefasst lässt sich formulieren: Eine unterstützende engagierte Haltung von Seiten der Schulleitung verleiht den traditionell sowohl in der Stundentafel als auch in der Auffassung vieler Akteure eher unterrepräsentierten Fächern Musik, Kunst und Darstellendes Spiel durch die angesprochenen Handlungsmöglichkeiten mehr Gewicht. Denn nur dann kann auch der musisch-kulturelle Bereich im Hinblick auf Ausgestaltung von Veranstaltungen und Außendarstellung der Schule qualitativ hochwertig aktiv werden. Ein gelingendes Zusammenspiel dieser Aspekte wird vielfältig positive Auswirkungen haben.

Lehr- und Lernformen sowie ergänzende Veranstaltungen

Im Kontext der Bläser- oder Chorklassenmodelle waren einander ergänzende Lernsettings von einerseits Einzel- und Kleingruppenunterricht und andererseits Gesamtgruppenunterricht anzutreffen. Diesbezüglich werden Formen des Team- bzw. Co-Teaching verfolgt: Meist übernimmt dabei der Fachlehrer die Proben mit dem Klassenorchester bzw. Klassenchor, während die Kooperationspartner, in dem Falle Instrumental- oder Gesangslehrer, für Übung und Unterricht in Kleingruppen zuständig sind.

Bewährte und / oder innovative Lehr- und Lernformen begegnen den Akteuren im alltäglichen Unterrichtsgeschehen. Dagegen ist der Bereich der *Veranstaltungen* an bestimmte epochale Ereignisse gebunden und im Vergleich zu anderen Fachbereichen vor allem ein grundlegender Bestandteil musisch-kultureller Praxen. Basierend auf den MUKUS-Daten werden Veranstaltungen im Sinne von *Ausstellungen* (betrifft den künstlerischen Bereich), *Aufführungen und Konzerten* (v. a. in den Bereichen Musik und Darstellendes Spiel), *öffentlichen Präsentationen* (damit sind meist Verbundaktionen gemeint) und anderen *Produktionen* (häufig betrifft das die Herstellung von fachspezifischen Produkten) verstanden. Das Zusammenspiel von Lehr- und Lernformen sowie ergänzenden Veranstaltungen gilt als Charakteristikum musisch-kultureller und insbesondere musikalischer Praxen. Es ermöglicht und erfordert das Eigenengagement und die Partizipation aller am Schulleben beteiligten Akteure. Somit kann festgehalten werden, dass der musisch-kulturelle Bereich mit seinen besonderen praxisnahen Inhalten und Konzepten, mit der Möglichkeit und der Bedingung besonderer Lehr- und Lernformen und zusätzlicher Veranstaltungen, die meist auf Ergebnispräsentationen innerhalb und

außerhalb des schulischen Systems zielen, alle beteiligten Personen in einen fachbezogenen und miteinander handelnden Zusammenhang bringt. Dieser besondere Zusammenhang zeichnet verantwortlich für vielfältige positive Befunde.

5 Ergebniskomplex II: Falldarstellung am Beispiel des Lernertrags bei Schülerinnen und Schülern

5.1 Der Lernertrag bei Schülerinnen und Schülern

Der Lernertrag für Schülerinnen und Schüler im Zusammenhang mit musikalischen Praxen betrifft drei unterschiedliche Bereiche. Im Rahmen von MUKUS konnten bereits bestehende Erkenntnisse zum Lernertrag bestätigt sowie fachspezifisch erweitert werden.¹¹ Demzufolge muss der Lernertrag dimensionalisiert betrachtet werden, und zwar hinsichtlich fachlicher, sozialer und persönlichkeitsbezogener Aspekte, jedoch sind alle drei Dimensionen nur selten losgelöst von einander zu betrachten. Meist stehen sie in einem engen, sich wechselseitig bedingenden Zusammenhang. Im Verlauf des Thematischen Kodierens sind drei Große Themen hervor gegangen, die wie folgt definiert worden sind:

„Der fachliche Lernertrag bezieht sich auf fachbezogene Informationen. Im Zentrum stehen Aussagen zu Fachkenntnissen, wie bspw. das Spielen eines Instrumentes, theoretische Kenntnisse oder andere meist praxisbezogene Lerninhalte.“

11 In der „Studie zur Entwicklung von Ganztagschulen“ (StEG) sind im Hinblick auf die Einschätzung des Ganztagschulnutzens aus Schülerperspektive (v. a. die Angebote betreffend) verschiedene Skalen zum Einsatz gekommen. Dabei wurden zwei zentrale Dimensionen betrachtet: „Lernförderlichkeit der Angebote sollte dann gegeben bzw. erfüllt sein, wenn die Schülerinnen und Schüler das Gefühl haben, der Besuch der Angebote helfe ihnen beim Lernen im Unterricht und trage zur Verbesserung der schulischen Leistungen bzw. Noten bei. Sozialer Nutzen hingegen sollte dann gegeben sein, wenn die Angebote Spaß und Freude bereiten und den Schülerinnen und Schülern erweiterte Möglichkeiten der sozialen Integration – etwa durch neue Freundschaften – eröffnen und darüber hinaus (unter Umständen) bestehende Betreuungsengpässe in der Familie beheben. (Radisch et al. 2006, S. 252)

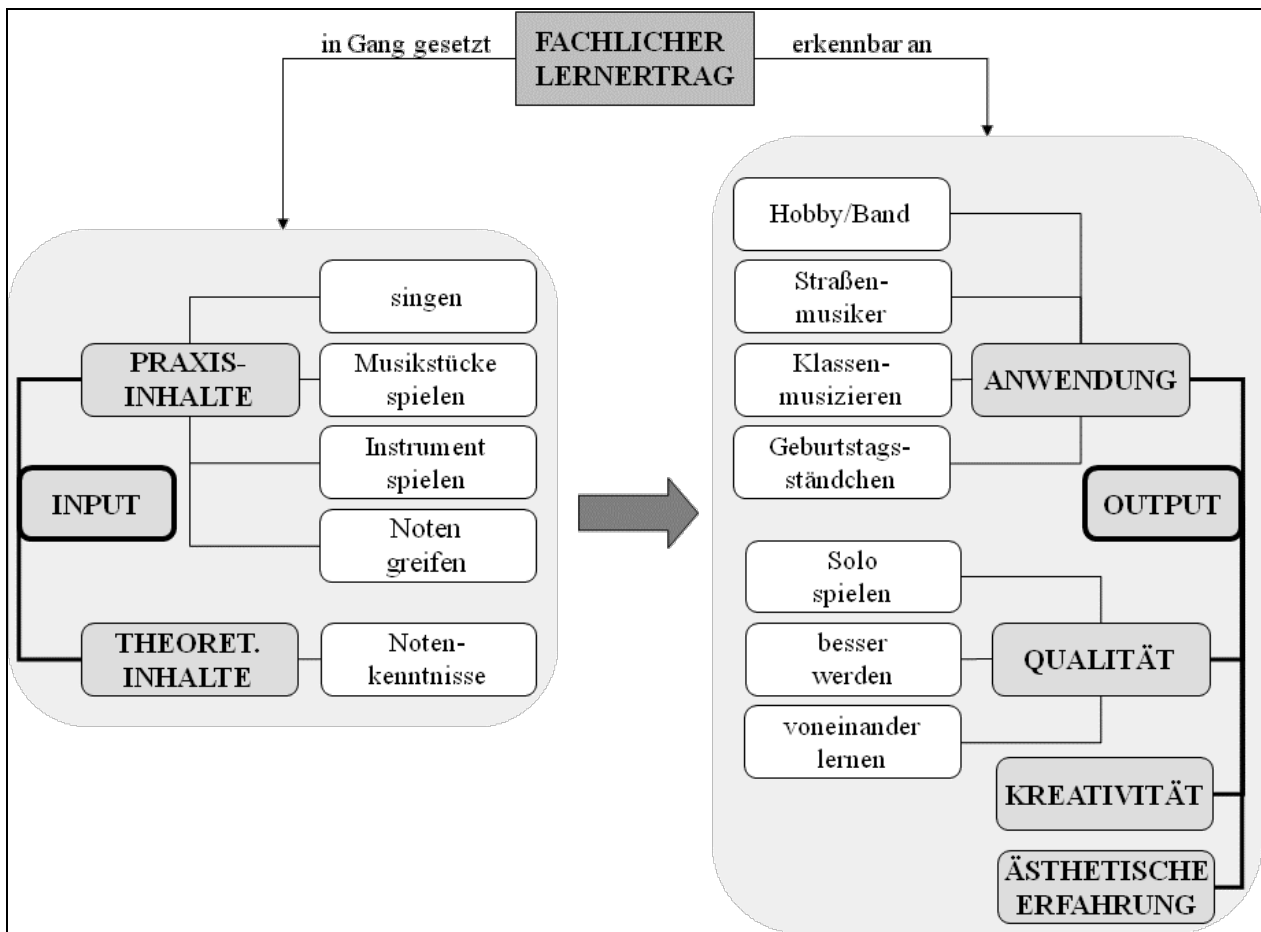


Abb. 5: Fachlicher Lernertrag

Quelle: eigene Darstellung MUKUS 2009 (Susanne Naacke)

„Der soziale Lernertrag steht als Oberbegriff für Informationen, die Gruppenprozesse und Interaktionen im Rahmen von musisch-kulturellen Lerninhalten und Angeboten thematisieren. Bspw. werden Aussagen zum Gruppengefühl, zu sozialen Kompetenzen, zu Gemeinschaftsaspekten oder zur Wahrnehmung der Einzelperson im Gruppengefüge thematisiert. Außerdem äußern sich die befragten Akteure im Zusammenhang mit musisch-kulturellen Praxen vermehrt zu ‚etwas macht mir Spaß‘.“

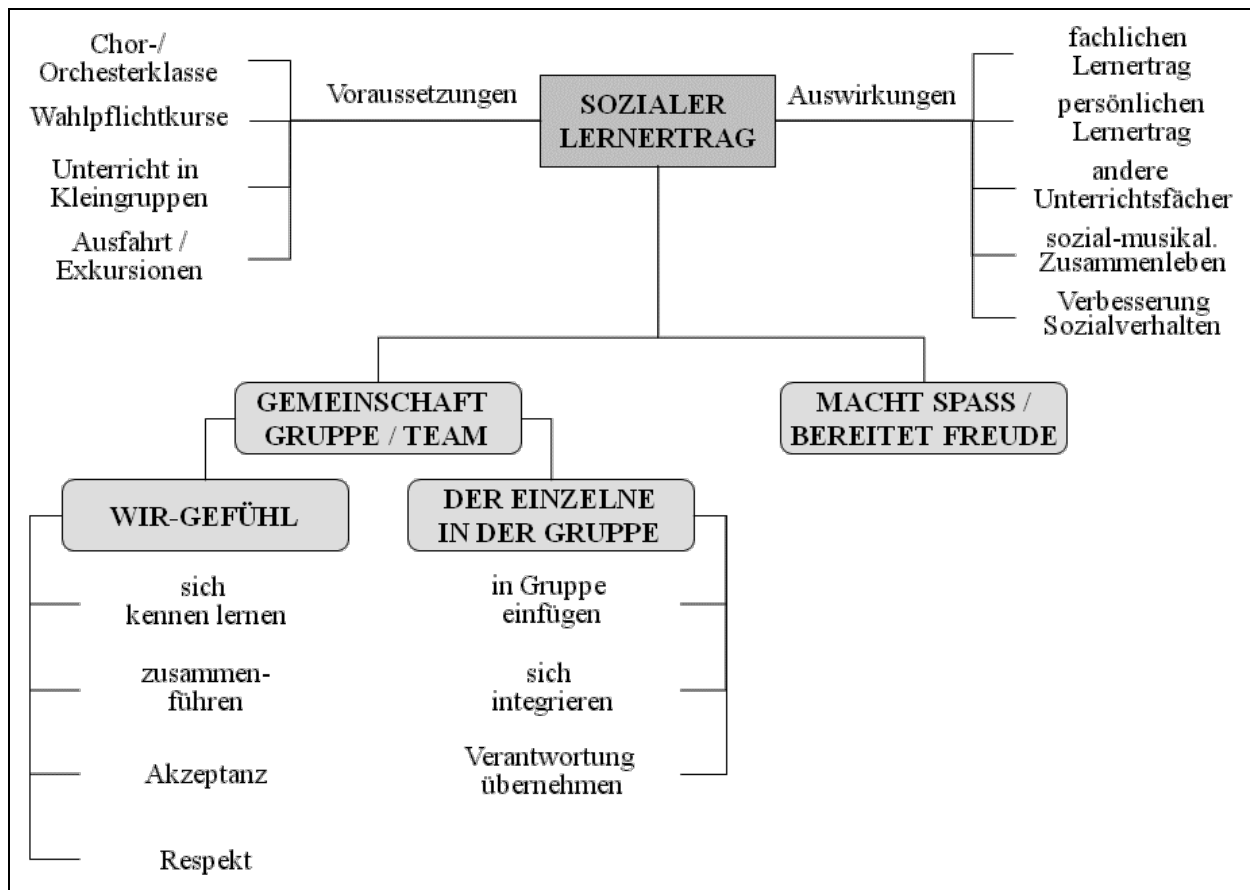


Abb. 6: Sozialer Lernertrag

Quelle: eigene Darstellung MUKUS 2009 (Susanne Naacke)

„Der persönlichkeitsbezogene Lernertrag im Zusammenhang mit dem musisch-kulturellen Bereich bezieht sich auf Informationen, die die Persönlichkeitsbildung des einzelnen Schülers thematisieren. Dabei stehen v. a. Aussagen zum Selbstwertgefühl, zum Selbstkonzept, Selbstvertrauen und -bewusstsein und ähnliches im Zentrum.“

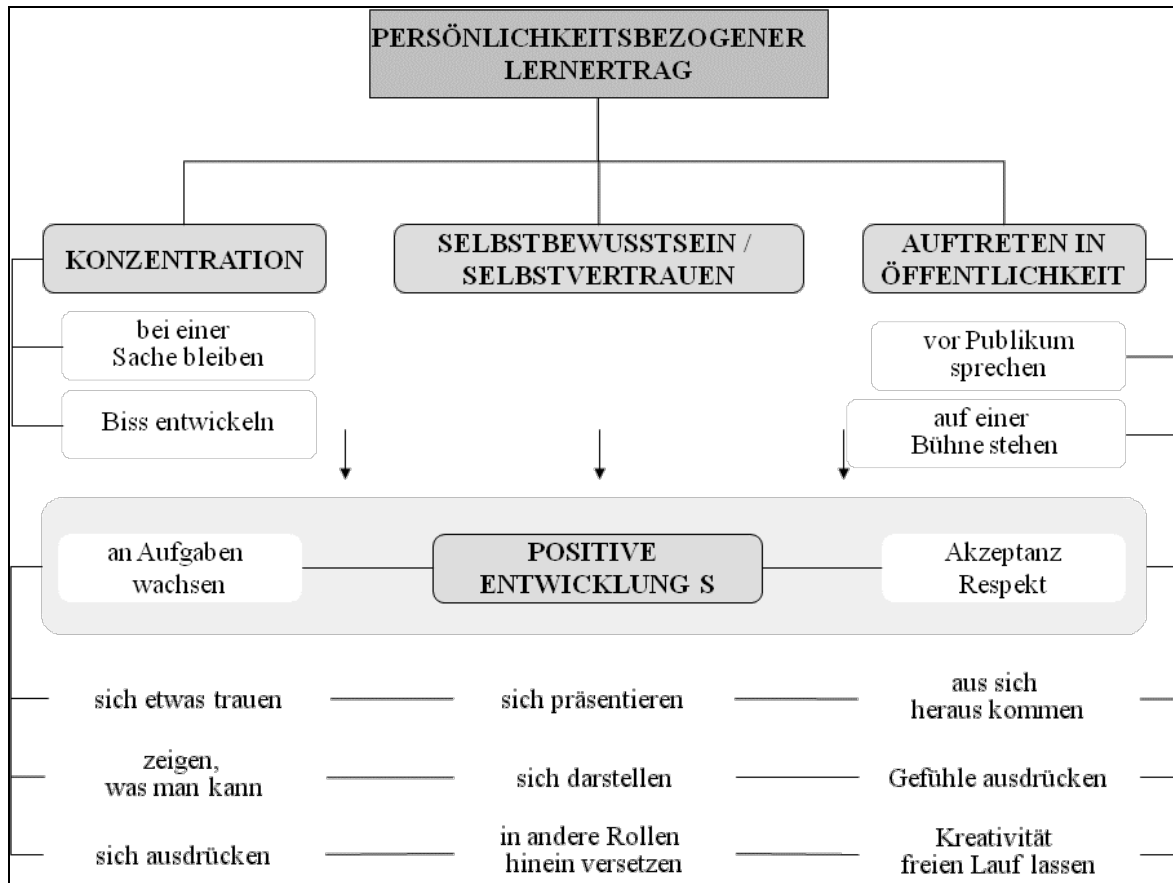


Abb. 7: Persönlichkeitsbezogener Lernertrag

Quelle: eigene Darstellung MUKUS 2009 (Susanne Naacke)

Folgende Punkte sind den Teildimensionen des Lernertrags gemeinsam: die Voraussetzungen oder Ursachen dafür liegen laut Aussagen der Interviewpartner in den charakteristischen Strukturmerkmalen musikalischer Praxen. Diese finden sich im schulischen Erklärungsmodell im Bereich der internen Strukturbedingungen. Dabei spielen die Praxisbezogenheit, die Verknüpfung von Unterricht und Ganztagsangeboten - also der Wechsel von Groß- und Kleingruppe - sowie die Stärkung der künstlerischen Fächer durch besondere Konzepte (Bläser-, Chorklasse) oder durch eine stundenmäßige Aufstockung eine bedeutsame Rolle. Aber allein das Vorhandensein sowie die Teilnahme am jeweiligen Unterricht ist nicht die Garantie für einen positiven Lernertrag bei den Schülerinnen und Schülern. Hier nehmen die Handlungs- und Interaktionsstrategien eine wesentliche Position ein. Grundlegend und charakteristisch ist dabei das erläuterte Interaktionsdreieck. Die bereichsspezifischen Lerninhalte werden mit bestimmten Lehr- und Lernformen umgesetzt, bspw. Unterricht in Kleingruppen, die einen verstärkten Einbezug der Akteure erfordern und ermöglichen. Zusätzliche Veranstaltungen sind oft Ziel vielfältiger Produktionen und bieten eine gute Plattform der Ergebnispräsentationen in

mehr oder weniger großer Öffentlichkeit. Dieses Zusammenspiel kann positiv ergänzt und verstärkt werden durch gelingende Kooperationsstrukturen, die weitere Lehr- und Lernformen ermöglichen.

Die Teildimensionen des Lernertrags werden wie folgt ersichtlich: der fachliche Lernertrag wird v. a. bei den Präsentationsformen im Unterricht oder bei den zusätzlichen Veranstaltungen deutlich. Der soziale Lernertrag kommt in den notwendigen Gruppenprozessen zum Tragen. Dabei steht dieser eng mit dem fachlichen Lernertrag in Verbindung. Werden Befunde aus der MUKUS-Studie einbezogen, so wird ersichtlich, dass gerade die musisch-kulturellen Praxen meist nur im Gruppenkontext gelingen. Jeder Einzelne ist bei Orchester- und Chorklassen, bei Theater- oder Musicalaufführungen, bei der Herstellung von Bühnenbildern oder gemeinsamen Ausstellungen wichtig. Nur dann können auch fachbezogene Anforderungen angemessen umgesetzt werden. Der persönlichkeitsbezogene Lernertrag wird am einzelnen Schüler deutlich. Er ist ebenso mit dem fachlichen Lernertrag verbunden, wird bei den benannten Präsentationsformen geformt und bildet sich entsprechend auch in diesem Rahmen ab.

Im folgenden Abschnitt wird aufgezeigt, inwieweit Schulkultur als Interpretations- und Orientierungsrahmen auf die Gestaltung von Lernkultur einwirkt und diesbezüglich die Entwicklung der drei Einzeldimensionen des Lernertrags unterstützt. Die Verbindung zwischen Schul- und Lernkultur wird insofern ersichtlich, als dass Schulkultur i. S. eines Ermöglichungsrahmens erst eine spezifische Lernkultur anregt und diese wiederum positiv auf die Ausgestaltung der Schulkultur rückwirkt. Das heißt, die eingangs formulierte Wirkungsweise von Organisationskultur auf Variablen wie Leistung oder Effektivität wird hier auf Individualebene – auf Schülerseite – anhand des Lernertrags ersichtlich. Anhand der Entwicklung des Gemeinschaftsgedankens am Domgymnasium kann das Zusammenwirken weiterer Faktoren illustriert werden.

5.2 Der Einfluss der Schulkultur auf die Dimensionen des Lernertrages am evangelischen Domgymnasium

Wenn Schulkultur Einfluss auf die Entwicklung einer für die beteiligten Akteure positiven und gelingenden Lernkultur hat, so muss gefragt werden, welchen Beitrag diesbezüglich die evangelische Profilbildung sowie die Ganztagschulkonzeption auf den *fachlichen Lernertrag* im Rahmen musikalischer Praxen leisten. Die intensive Durchführung von Chor- und Bläserklassen ist

zunächst an die Ermöglichung einer zusätzlichen Stunde Musik gebunden, die aufgrund der gebundenen Ganztagsform zur Verfügung gestellt werden kann. Auch zusätzliche Angebote wie etwa der „Freiwilligenchor“ werden aufgrund dieser Angebotsstruktur ermöglicht. Ebenso lässt sich ein unterstützender Einfluss der evangelischen Orientierung ausmachen. Durch die intensive Anbindung an den Dom zu Brandenburg ist die Schule aufgefordert, Gottesdienste oder andere kirchliche Feiern auszugestalten. Diese Partnerschaft ermöglicht einen Großteil der zusätzlichen Veranstaltungen und Präsentationsmöglichkeiten, die für das zentrale Interaktionsdreieck so bedeutsam sind. Wenn der Schulleiter wie folgt formuliert: „Das Religiöse und das Musische, die sind ganz nah beieinander“, so wird deutlich, dass die Säulen der Schulkultur mit Bedacht ausgewählt worden sind und in ihrer ergänzenden Zusammenführung jene Lernkultur unterstützen, die positive fachliche Lernerträge hervorbringt.

Bezugnehmend auf die Wechselwirkung zwischen Schul- und Lernkultur hinsichtlich *persönlichkeitsbezogener Lernerträge*, muss die Verzahnung mit der evangelischen Orientierung hervorgehoben werden. Denn auch das Religiöse ist wie die Musik eine zusätzliche Ebene, die dabei hilft, subjektive Erfahrungen zu machen und zu reflektieren. Wenn der Schulleiter formuliert, dass das Religiöse und das Musische sehr nah beieinander sind, so begründet er diese elementare Aussage mit den Worten:

„Es geht um emotionale Bildung. Es geht darum, dass man empfindsam wird. Und dass man das ausdrücken lernt. Sprachlich, gestalterisch, musikalisch, wie auch immer. Und.. insofern hat das viel mit Menschenbildung im Allgemeinen zu tun.“ (Schulleiter)

Auch der *soziale Lernertrag* impliziert Facetten, die über die Musik hinausgehen. Allerdings geht es nun um individuelle Erfahrungen im Gruppenkontext. Des Weiteren finden hier Aussagen zu „Spaß und Freude“ ihren Ausdruck. Dass die musisch-kulturellen und insbesondere musikalischen Praxen ein Gemeinschaftserlebnis unterstützen, ist nicht unbekannt. Dennoch ist es bemerkenswert, welche Rolle die Wahrnehmung dieser Gemeinschaftserfahrung für die befragten Akteure hat. Gerade jene Aspekte werden meist an erster Stelle genannt. Es liegt auf der Hand, dass vor allem bei musikalischen Praxen die Gemeinschaft unerlässlich für den Erfolg der Umsetzung und Darbietung ist. Ein Orchesterstück klingt erst dann gut, wenn wirklich alle dafür vorgesehenen Instrumente mitspielen. Ein Chor ist erst dann ein Chor, wenn verschiedene Stimmgruppen in ausreichendem Umfang vorhanden sind und die Harmonien zum Klingen bringen.

„Sie müssen ja zusammen spielen. Es geht auch schon damit los, wenn Sie die äh.. Chorklasse haben. Einer singt immer verkehrt. Passt nicht rein! Ich muss in einer Gemeinschaft zusammen spielen. Die Bläserklasse genauso. .. Ich muss! Und dadurch lerne ich. Ich bin zwar selbst immer noch derjenige, der spielt, aber ich muss mich integrieren.“ (Mutter, Elternrat)

Der Gemeinschaftsgedanke ist außerdem grundlegender Bestandteil sowohl der evangelischen Profilorientierung als auch ganztägiger Formen und soll nachfolgend erläutert werden.

Die Bedeutung der Gemeinschaft am Domgymnasium

Gemeinschaft und Gemeinde sind zwei Begriffe, die etymologisch und inhaltlich sehr nah beieinander liegen. Wie in der Schulkonzeption dargelegt, verstehen sich die am schulischen Leben beteiligten Akteure des Domgymnasiums als eine schulische Gemeinde.

„Ja, die Gemeinschaft würde ich schon sagen. Eben, dass es eine kirchliche Schule ist.“ (Schüler)

Die evangelische Profilorientierung setzt bewusst auf die Ausgestaltung eines gemeinschaftlichen Lebens. Dabei werden vielfältige Elemente implementiert, die diesen Gemeinschaftsgedanken unterstützen. Am Domgymnasium ist es zum einen das gemeinsame Mittagessen: Die befragten Schülerinnen und Schüler formulieren, dass sie es als sehr schön empfinden, mit den Lehrerinnen und Lehrern beim Mittagessen auch mal „ganz normal“ zu reden.

„Gerade beim Mittagessen. Die Lehrer essen meistens mit und wir sitzen an Tischen. [...] Das ist ganz schön. Und da ist wirklich ein intensives Gespräch zwischen Lehrern und Schülern während des Mittagessens.“ (Schülerin)

Auch die obligatorisch am Wochenbeginn stattfindende Klassenratsstunde unterstützt den Gemeinschaftsgedanken. Darin können Probleme oder erfolgreiche Momente, die im Rahmen der Klassengemeinschaft aufgetreten sind, thematisiert werden.

Zusätzlich dazu ist auch der Ganztagsschulkonzeption die Idee der Gemeinschaft inhärent. Dieser Umstand liegt in ihrer Entstehungsgeschichte begründet. Stellt die Ganztagsschule nicht nur einen Lernort sondern zudem einen Lebensraum dar, so ermöglicht sie den Kindern und Jugendlichen, die in

einem dünn besiedelten Bundesland wie Brandenburg aus einem großen Einzugsbereich anreisen, dass sie viel Zeit miteinander verbringen und ihre Freunde treffen können.

„Dass man auch seine Freunde die ganze Zeit sieht, weil wir kommen ja wirklich von Ecken [...] aus Brandenburg. Von daher sieht man wirklich die Freunde immer wieder.“ (Schülerin)

Zudem ist es aufgrund der vermehrt auftretenden Ein-Kind-Familien notwendig, dass die Schülerinnen und Schüler in alternativen sozialen Gruppen Gemeinschaft mit anderen erfahren. Mit jahrgangsübergreifenden oder fächerverbindenden Lernarrangements unterstützt die Ganztagschulkonzeption ein gegenseitiges Kennenlernen sowie gemeinschaftliches Lernen im Rahmen der Schule.

„Es ist einfach, weil wir alle viel miteinander zu tun haben, v. a. jetzt, wenn wir noch nicht so viele sind.“ (Schüler)

„It's English Time!“ So wurde unerwartet die Mittagspause am Donnerstag von den Lehrkräften eingeläutet. Während dieser Zeit haben ausnahmslos alle schulischen Akteure Englisch miteinander gesprochen: Ohne Frage hat „It's English Time“ einen Spaßcharakter, der einmal pro Woche die schulische Situation auf besondere Weise auflockert und den Gemeinschaftsgedanken fördert.

Dass zusätzlich dazu der Schulleiter alle Schülerinnen und Schüler zu Schuljahresbeginn mit Namen kennt, dass der befragte Kooperationspartner der Bläserklasse sich eindeutig nicht als *Partner*, sondern als *Mitglied* des Lehrerkollegiums fühlt, dass die Schülerinnen und Schüler wissen, sie können ihren Fragen und Anregungen, ihren individuellen Problemen jederzeit Gehör verschaffen, weil ein vertrauensvolles Verhältnis zwischen den beteiligten Akteuren vorherrscht – all das sind neben den genannten Schwerpunkten kleinere aber nicht weniger bedeutende Elemente, die zur Formung und Wahrnehmung einer guten Gemeinschaft beitragen.

Insgesamt formuliert eine Schülerin der Schülervvertretung Folgendes auf die Frage nach der besonderen Stärke der Schule:

„Na ganz definitiv das Zwischenmenschliche. Das wird hier stark hervorgehoben, weil wir halt dieses Ganztagskonzept haben und

dadurch viel Zeit haben uns auch intensiver mit den Persönlichkeiten zu beschäftigen.“ (Schülerin)

Die Gestaltung der Gemeinschaft findet nicht nur in den musikalischen Praxen statt, sondern wird, wie dargelegt, auch durch die Ganztagschulkonzeption sowie die evangelische Profilorientierung maßgeblich angeregt. Es sind die drei Schulkultursäulen im Verbund, die zu einer vertrauensvollen Gemeinschaft als Grundlage für ein besseres Lehren und Lernen beitragen.

Zusammenfassend lassen sich die Einzeldimensionen des Lernertrags am Domgymnasium wie in *Abbildung 8* darstellen. Dabei wird zudem der positive Einfluss der zwei Schwerpunkte der Schulkultur – Ganztagskonzeption sowie evangelische Profilorientierung – verdeutlicht.

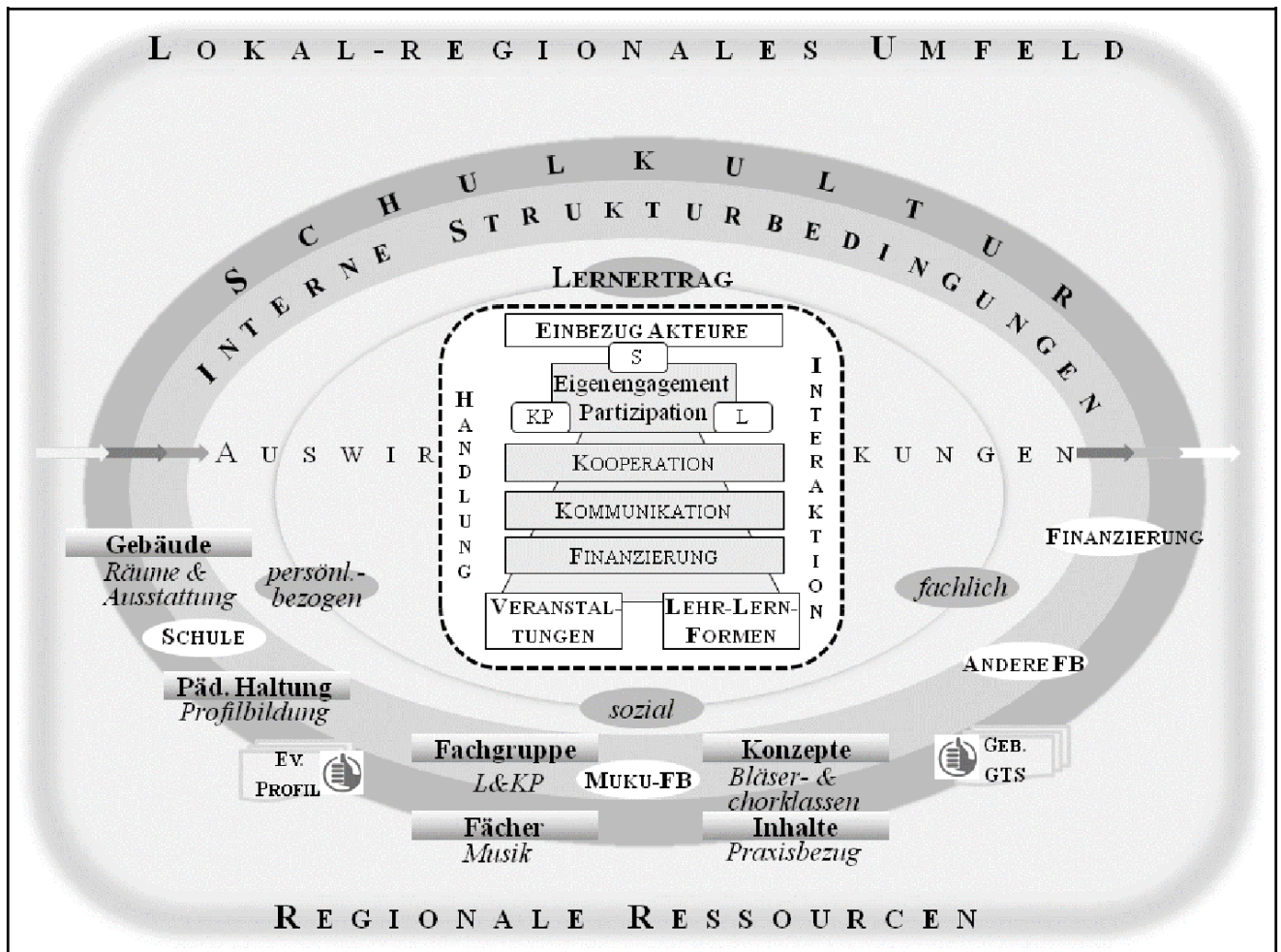


Abb. 8: Dimensionen des Lernertrags im Fach Musik im Schulmodell

Anmerkungen: vgl. Abbildung 3; S = Schülerinnen und Schüler, L = Lehrerinnen und Lehrer, KP = Kooperationspartner;

👍 = hat positiven Einfluss; 👎 = hat negativen Einfluss

Quelle: eigene Darstellung 2010 (Susanne Naacke)

6 Theorie der Schulentwicklung

Im Folgenden soll ein Merksatz zur Schulentwicklungstheorie vorgestellt werden, der die angesprochenen Dimensionen umfasst. Wenn Unterrichtsentwicklung als Kern von Schulentwicklung gilt, so wird im Merksatz in besonderer Weise auf das Interaktionsdreieck als Erklärungsschlüssel für die Lehr- und Lernarrangements von Chor- und Bläserklasse Bezug genommen.

Schulentwicklung setzt voraus:

... ein spezifisches Bildungsverständnis und die „Lernende Schule“

Gelingende Schulentwicklung beruft sich zum einen auf ein spezifisches zugrunde liegendes Bildungsverständnis und setzt zum anderen Schulen als „Lernende Organisationen“ voraus. „Lernende Schulen“ sind demnach in der Lage, im Spannungsfeld von gesetzten Standards und individuellem Entwicklungspotenzial zunehmend selbstgesteuert und selbstreflektiert auf die veränderten gesellschafts- und bildungspolitischen Herausforderungen zu reagieren.

Schulentwicklung fokussiert:

... Unterricht als Kernelement – neue Lehr- und Lernarrangements

Unterricht als Kernelement von Schule und Schulentwicklungsprozessen unterliegt ausgehend vom Bildungsverständnis grundlegenden Veränderungen mit dem Ziel der Herausbildung neuer Lehr-Lernarrangements.

→ Interaktionsdreieck musikalischer Praxen – der Erklärungsschlüssel für die neuen Lehr- und Lernarrangements und damit verbundener positiver Auswirkungen

Mit Chor- und Bläserklassen gehen neue Lehr- und Lernarrangements einher, die durch folgenden Dreischritt charakterisiert werden können: Demgemäß ist eine erfolgreiche Umsetzung gekennzeichnet durch erstens spezifische Lehr- und Lernformen, die aus einem Wechsel von Groß- und Kleingruppenunterricht bestehen, zweitens zusätzlichen (öffentlichen) Veranstaltungen oder Präsentationsformen (wie bspw. Konzerten). Dies ermöglicht drittens in besonderer Weise einen intensiven Einbezug sämtlicher Akteure (neben Schüler- und Lehrerschaft ist die Einbindung von Eltern, Schulleitung sowie anderen Beteiligten wie dem Schulförderverein unerlässlich für die Gestaltung von Chor- und Bläserklassen).

Dieser Dreischritt – *Interaktionsdreieck* genannt – kann durch vielfältige Kooperationsstrukturen, durch eine gute Kommunikation sowie durch solide Finanzierungsmodelle entscheidend unterstützt werden.

Schulentwicklung ist eingebettet in:

... Schulkultur und Lernkultur

Die neuen Lehr- und Lernarrangements sind in eine spezifische Schulkultur als Interpretations- und Orientierungsrahmen eingebettet. Indem sich Schul- und Lernkultur gegenseitig bedingen, werden vielfältige Wechselwirkungen zwischen den beteiligten Akteuren im Rahmen der Lehr- und Lernsettings innerhalb der Schule sowie zwischen Schule und Umwelt angestoßen.

Schulentwicklung hat zum Ziel:

... Qualitätsentfaltung und -sicherung der Einzelschule (Schulentwicklung erster Ordnung)

Ziel ist stets die Qualitätsentfaltung sowie -sicherung zunächst der Einzelschule, die Lernprozesse auf Individual-, Gruppen- und Organisationsebene von unterschiedlicher Komplexität ermöglicht.

... Schulentwicklung zweiter und dritter Ordnung

In Anlehnung an das Drei-Wege-Modell nach Kempfert & Rolff (1999) geht es im Folgenden auch um die Etablierung „Lernender Schulen“ sowie um die Entwicklung von übergeordneten Koordinations- und Unterstützungssystemen.

6.1 Schulspezifische Erläuterung für das Domgymnasium

Zur Unterrichts-, Personal- und Schulentwicklung

Am Beispiel des Domgymnasiums entfalten sich nicht nur, wie bereits dargelegt, alle drei Formen der Schulentwicklung nach Kempfert & Rolff (1999) (vgl. *Abschnitt 2.2*). Im Zusammenhang mit dem Mehrebenenmodell der Schulentwicklung nach Wiater (2002) sind schulische Entwicklungsprozesse zudem auf Unterrichts-, Personal- sowie Organisationsebene zu verzeichnen. In diesem Verständnis gilt Unterricht als Kernelement von Schulentwicklung einer Einzelschule. Die Durchführung der Lehr- und Lernarrangements von Chor- und Bläserklasse und die damit einhergehenden positiven Befunde sind zusätzlich zu den musikspezifischen Merkmalen an notwendige Strukturbedingungen, wie bspw. zusätzliche Unterrichtszeit, eine Vielzahl an Räumlichkeiten oder die Möglichkeit von Aufführungsorten und -gelegenheiten gebunden. Diesbezüglich ergänzen sich die Schwerpunkte der Schulkultur vielfältig. Somit schaffen die evangelische Profilorientierung und die Ganztagschulkonzeption günstige, zusätzlich unterstützende Rahmenbedingungen für

die Umsetzung musikalischer Praxen. Veränderte Lernsettings sind ausgehend von der Betrachtung musikalischer Praxen am Domgymnasium auch darüber hinaus grundlegender Bestandteil der Unterrichtsgestaltung und -entwicklung.

Insgesamt kann gesagt werden, dass das Interaktionsdreieck in den Aussagen der befragten der Personen *der* Erklärungsschlüssel ist für positive Befunde auf Individualebene (positiver Lernertrag auf Schülerseite, hohe Motivation und hohe Bereitschaft zu Innovation auf Lehrerseite), auf Gruppenebene (positive Gestaltung und Wahrnehmung von Schul- und Unterrichtsklima) sowie auf Schulebene (eine erfolgreiche Öffnung der Schule nach innen und außen). Aufgrund ihrer Eigenschaft als „Lernende Schule“ kann das Domgymnasium selbstständig weiterführende Perspektiven, Kurskorrekturen und notwendige Entwicklungsschritte vollziehen.

7 „Aber das Musikalische, das ist das A und O!“ – Schulspezifischer Ausblick

Zusammenfassend kann gesagt werden, dass Schulentwicklungsprozesse am Domgymnasium nicht abgeschlossen sind, sondern mit einem hohen Verantwortungsgefühl stetig voran getrieben werden müssen: Chor- und Bläserklassen werden sich nicht auf die Jahrgänge sieben und acht beschränken, sondern möglicherweise eine Weiterführung in Schulchor oder -orchester sowie in grundlegenden konzeptionellen Überlegungen erfahren. Mit der Erweiterung auf die gymnasiale Oberstufe werden zusätzliche Herausforderungen einhergehen, denen sich die Schule mit Bedacht annehmen wird: Die gewünschte Vergrößerung der Schüler- und Lehrerschaft wird ebenso zu bewältigen sein, wie die mit der Sekundarstufe verbundenen zusätzlichen Lehr- und Lernarrangements, bspw. das „Diakonische Praktikum“. Es bleibt zu wünschen, dass sämtliche Überlegungen stets ihre Verankerung in der Schulkultur und im gemeinsamen Dialog finden werden. Wenn also „das Musikalische das A und O [ist]“, wie eine Mutter formulierte, so verdeutlicht das den Stellenwert der Chor- und Bläserklassen in der Perspektive der beteiligten Personen.

„Aber der Donnerstag, das ist ein richtig klingender Tag. Weil dann, wenn die Fenster auf sind, dann tutet das hier über die ganze Dominsel. Das freut mein Herz.“ (Schulleiter)

Die Untersuchung am Domgymnasium hat gezeigt, dass gebundene Ganztagschulform und evangelische Trägerschaft kein schulpolitisches Korsett sein müssen, sondern – wenn sie sorgfältig und in Abstimmung mit der Um-

welt gestaltet werden – einen anregenden Rahmen vorgeben können. Eine so verstandene Schulkultur als Interpretations- und Verständigungsrahmen wird Schulentwicklungsprozesse im Spannungsfeld von staatlicher Steuerung und zunehmender Eigenverantwortlichkeit der Einzelschule konstruktiv unterstützen.

Literatur

Argyris, Chris; Schön, Donald A.; Rhiel, Wolfgang (1999): Die lernende Organisation. Grundlagen, Methode, Praxis. Stuttgart: Klett-Cotta (/Management - Die blaue Reihe]).

Bildungskommission NRW (1995): Zukunft der Bildung - Schule der Zukunft. Neuwied: Luchterhand.

Bogner, Alexander; Menz, Wolfgang (2005): Das theoriegenerierende Experteninterview. Erkenntnisinteresse, Wissensformen, Interaktion. In: Bogner, Alexander; Littig, Beate; Menz, Wolfgang (Hg.): Das Experteninterview. Theorie, Methode, Anwendung. 2. Aufl. Wiesbaden: VS Verl. für Sozialwiss., S. 33- 70.

Creswell, John W. (2005): Qualitative inquiry and research design. Choosing among five traditions. Thousand Oaks Calif.: Sage.

Eichler, Dirk (2008): Veränderungsprozesse pädagogischer Institutionen. Organisationstheoretische Reflexionen vor dem Hintergrund gestiegener Anforderungen. Wiesbaden: VS Verlag für Sozialwissenschaften / GWV Fachverlage GmbH Wiesbaden (Springer-11776 /Dig. Serial]).

Flick, Uwe (2007): Qualitative Sozialforschung. Eine Einführung. Orig.-Ausg., vollst. überarb. und erw. Neuausg. Reinbek bei Hamburg: Rowohlt-Taschenbuch-Verl. (Rororo Rowohlts Enzyklopädie, 55694).

Flick, Uwe (2008): Triangulation. Eine Einführung. 2. Aufl. Wiesbaden: VS Verl. für Sozialwiss. (Qualitative Sozialforschung, 12).

Flick, Uwe (r2007 /// 2008): Triangulation in der qualitativen Forschung. In: Flick, Uwe; Kardorff, Ernst von; Steinke, Ines (Hg.): Qualitative Forschung. Ein Handbuch. 6., durchges. und aktualisierte Aufl., Orig.-Ausg. Reinbek bei Hamburg: Rowohlt-Taschenbuch-Verl. (Rororo Rowohlts Enzyklopädie, 55628), S. 309-318.

Flick, Uwe; Kardorff, Ernst von; Steinke, Ines (Hg.) (r2007 /// 2008): Qualitative Forschung. Ein Handbuch. 6., durchges. und aktualisierte Aufl.,

- Orig.-Ausg. Reinbek bei Hamburg: Rowohlt-Taschenbuch-Verl. (Rororo Rowohlts Enzyklopädie, 55628).
- Franzpötter, Reiner (1997): Organisationskultur. Begriffsverständnis und Analyse aus interpretativ-soziologischer Sicht. Zugl.: Aachen, Techn. Hochsch., Diss., 1995. 1. Aufl. Baden-Baden: Nomos-Verl.- Ges. (Nomos-Universitätsschriften : Kulturwissenschaft).
- Fuchs, Max (2009): Kultur lernen - Auf dem Weg zu einer Kulturschule. In: Kulturelle Bildung, Jg. 3, H. 3, S. 4- 6.
- Kempfert, Guy; Rolff, Hans-Günter (1999): Pädagogische Qualitätsentwicklung. Ein Arbeitsbuch für Schule und Unterricht. Weinheim u.a.: Beltz (Neue Lehrerbildung und Schulentwicklung).
- Klieme, Eckhard; Avenarius, Hermann; Blum, Werner, et al. (Hg.) (2003): Zur Entwicklung nationaler Bildungsstandards. Eine Expertise. Berlin: BMBF (Bildungsforschung, 1).
- Criblez, Lucien; Oelkers, Jürgen; Reusser, Kurt; Berner, Esther; Halbheer, Ueli; Huber, Christina (2009): Bildungsstandards. Zug: Klett und Balmer (Lehren lernen - Basiswissen für die Lehrerinnen- und Lehrerbildung).
- Knigge, Jens; Lehmann-Wermser, Andreas (2008): Bildungsstandards für das Fach Musik: Eine Zwischenbilanz. In: Zeitschrift für Kritische Musikpädagogik, Sonderedition: Bildungsstandards und Kompetenzmodelle für das Fach Musik?, S. 60–98. Verfügbar unter: www.zfkm.org/sonder08-knigge-lehmannwermser.pdf [18.02.2010].
- Laging, Ralf; Hildebrandt-Stramann, Reiner; Teubner, Jürgen: STUBSS - Studie zur Entwicklung von Bewegung, Sport und Spiel in der Ganztagschule. Verfügbar unter: <http://www.unimarburg.de/fb21/ifsm/ganztagschule> [18.02.2010].
- Lamnek, Siegfried (2005): Qualitative Sozialforschung. Lehrbuch. 4., vollst. überarb. Aufl. Weinheim, Basel: Beltz.
- Lehmann-Wermser, Andreas; Naacke, Susanne & Nonte, Sonja (in Druck): Musisch-kulturelle Bildung an Ganztagschulen. Empirische Befunde, Chancen und Perspektiven. Weinheim: Juventa-Verl. (Studien zur ganztägigen Bildung).
- Merriam, Sharan B. (1988): Case study research in education. A qualitative approach. San Francisco: Jossey-Bass Publ. (Jossey-Bass education series).

- Metz-Göckel, Hellmuth (1996): Einstellungen und Werthaltungen in Organisationen. Essen: Verl. Die Blaue Eule (Psychologie zwischen Theorie und Anwendung, 6).
- Naacke, Susanne; Lehmann-Wermser, Andreas (2009): MUKUS - Studie zur musisch-kulturellen Bildung an Ganztagsschulen. Qualitative Fallstudien. In: Schläbitz, Norbert (Hg.): Interdisziplinarität als Herausforderung musikpädagogischer Forschung. Essen: Verl. Die Blaue Eule (Musikpädagogische Forschung, 30), S. 97–124.
- Niedersächsisches Landesamt für Lehrerbildung und Schulentwicklung (NiLS): SEIS Deutschland.
Verfügbar unter: <http://www.seis-deutschland.de/> [18.02.2010].
- Radisch, F., Stecher, L., Klieme, E., Kühnbach, O. (2007): Unterrichts- und Angebotsqualität aus Schülersicht. In: Holtappels, H. G., Klieme, E., Rauschenbach, T., Stecher, L. (Hrsg.): Ganztagsschule in Deutschland. Ergebnisse der Ausgangserhebung der „Studie zur Entwicklung von Ganztagsschulen“ (StEG). S. 227-260, Weinheim: Juventa-Verl.
- Rahm, Sibylle (2005): Einführung in die Theorie der Schulentwicklung. Weinheim: Beltz (Beltz Studium: Erziehung und Bildung).
- Reinders, Heinz (2005): Qualitative Interviews mit Jugendlichen führen. Ein Leitfaden. München: Oldenbourg.
- Risse, Erika (2003): Lernkultur als Ziel einer systemischen Entwicklung der Schule. Oldenburg: Bibliotheks- und Informationssystem der Univ.
- Rosenstiel von, Lutz (2007 /// 2008): Organisationsanalyse. In: Flick, Uwe; Kardorff, Ernst von; Steinke, Ines (Hg.): Qualitative Forschung. Ein Handbuch. 6., durchges. und aktualisierte Aufl., Orig.-Ausg. Reinbek bei Hamburg: Rowohlt-Taschenbuch-Verl. (Rororo Rowohlts Enzyklopädie, 55628), S. 224–238.
- Schein, Edgar H. (1995): Unternehmenskultur. Ein Handbuch für Führungskräfte. Frankfurt/Main: Campus-Verl.
- Schreyögg, Georg (2008): Organisation. Grundlagen moderner Organisationsgestaltung ; mit Fallstudien. 5., vollst. überarb. und erw. Aufl. Wiesbaden: Gabler (Lehrbuch).
- Senge, Peter M (1998): Die fünfte Disziplin. Kunst und Praxis der lernenden Organisation. Stuttgart: SchäfferPoeschel.

- Stake, Robert E. (2000): Case Studies. In: Denzin, Norman K.; Lincoln, Yvonna S. (Hg.): Strategies of qualitative inquiry. [Nachdr.]. Thousand Oaks: SAGE Publ. (Handbook of qualitative research / Norman K. Denzin Yvonna S. Lincoln eds., Vol. 2), S. 86-109.
- Strauss, Anselm; Corbin, Juliet (1996): Grounded theory. Grundlagen qualitativer Sozialforschung. Weinheim: Beltz PsychologieVerlagsUnion.
- Strauss, Anselm L.; Hildenbrand, Astrid (1998): Grundlagen qualitativer Sozialforschung. Datenanalyse und Theoriebildung in der empirischen soziologischen Forschung. 2. Aufl. München: Fink (UTB für Wissenschaft Uni-Taschenbücher, 1776, Soziologie).
- Strübing, Jörg (2008): Grounded theory. Zur sozialtheoretischen und epistemologischen Fundierung des Verfahrens der empirisch begründeten Theoriebildung. 2., überarb. und erw. Aufl. Wiesbaden: VS Verl. für Sozialwiss. (Qualitative Sozialforschung, 15).
- Wiater, Werner (2002): Theorie der Schule. 1. Aufl. Donauwörth: Auer (Prüfungswissen - Basiswissen - Schulpädagogik).
- Yin, Robert K. (2005): Case study research. Design and methods. 3. ed., [Nachdr.]. Thousand Oaks, Calif.: Sage (Applied social research methods series, 5).

Erfassung der Kompetenz „Musik wahrnehmen und kontextualisieren“ mit Hilfe von Testaufgaben

Aufgabenentwicklung und -analyse im Projekt KoMus

“Historically, task design has been regarded more as an art than a science.”

(Mislevy, Steinberg & Almond 2002, S. 98)

1 Einleitung

„Kompetenzorientierung“ ist zum pädagogischen und bildungspolitischen Modebegriff avanciert: Schulen entwickeln Programme zur Förderung von Methoden- und Sozialkompetenz; nationale Bildungsstandards reagieren auf die PISA-Misere mit der Formulierung von zu erreichenden Kompetenzen, und in der psychologischen und allgemein-pädagogischen Literatur ist die Anzahl der Veröffentlichungen zum Thema Kompetenzen in den letzten 10 Jahren ‚explodiert‘ (Klieme & Hartig 2007, S. 13). Im Fach Musik wird die Kompetenzorientierung auf curricularer Ebene breit – oft jedoch leider auch unreflektiert – implementiert (Knigge & Lehmann-Wermser 2008). Der fachdidaktische Diskurs gestaltet sich allerdings sehr heterogen: Partielle Zustimmung erhält das Kompetenzkonzept bspw. im Kontext des „aufbauenden Musikunterrichts“ (Jank 2007), bei anderen Autoren¹ trifft es auf Skepsis bis hin zu kategorischer Ablehnung (z. B. Richter 2008, 2009). Die Hintergründe für die unterschiedlichen Reaktionen zu beleuchten, würde eine eigene Publikation erfordern, aber ein Teil der konträren Einschätzungen hat sicherlich mit der Unsicherheit darüber zu tun, was der Kompetenzbegriff eigentlich bezeichnet: Die Kritiker finden ihn zu kognitionslastig und fürchten angesichts des Booms der Kompetenzmessungen eine Abwendung vom klassischen Bildungsbegriff; die Befürworter betonen gerade die Vieldimensionalität des Kompetenzkonzepts und meinen, dass es zur Steigerung von Unterrichtsqualität und durchaus

1 Aus Gründen der leichteren Lesbarkeit verzichten wir im Folgenden auf die Nennung beider Geschlechter.

auch zur Bildung beitragen kann, wenn Schüler ihre Kompetenzen entwickeln und ausbauen (Klieme & Hartig 2007,3 S. 22). Die Kompetenzdiskussion wird jedenfalls mit Leidenschaft geführt – auch in der Musikpädagogik (Knigge & Lehmann-Wermser 2008; Niessen 2009).

Die Unterschiedlichkeit der Einschätzung hängt sicherlich mit der Geschichte und den Verwendungszusammenhängen des Kompetenzbegriffs zusammen, der in Sprachwissenschaft, Psychologie und Erziehungswissenschaften jeweils sehr unterschiedliche Bestimmungen erfahren hat (für eine ausführliche Diskussion siehe z. B. Klieme & Hartig 2007). Klieme & Hartig (2007) fassen „zentrale Bestandteile des Begriffsverständnisses“ folgendermaßen zusammen: „Kompetenzen sind Dispositionen, die im Verlauf von Bildungs- und Erziehungsprozessen erworben (erlernt) werden und die Bewältigung von unterschiedlichen Aufgaben bzw. Lebenssituationen ermöglichen. Sie umfassen Wissen und kognitive Fähigkeiten, Komponenten der Selbstregulation und sozial-kommunikative Fähigkeiten wie auch motivationale Orientierungen. Pädagogisches Ziel der Vermittlung von Kompetenzen ist die Befähigung zu selbstständigem und selbstverantwortlichem Handeln und damit zur Mündigkeit“ (S. 21). Die Autoren betonen, dass ein solches Begriffsverständnis sehr wohl mit der bekannten Definition von Weinert konform geht, die u. a. das Kompetenzverständnis im Rahmen von Bildungsstandards bestimmt (Weinert 2001; vgl. Klieme et al. 2003). Pädagogisch geprägt ist die Beschreibung von Klieme & Hartig (2007, S. 21) u. a. durch die Betonung der Erlernbarkeit von Kompetenzen und die Zielbestimmung, die das Kompetenzkonzept als nicht nur deskriptives ausweist.

Ein solch facettenreiches und pädagogisch akzentuiertes Verständnis des Begriffs wurde auch der Forschung im Projekt *KoMus* zugrunde gelegt, aus dem im Folgenden einige Ergebnisse der Aufgabenentwicklung und -analyse dargestellt werden. Im Rahmen dieses von der DFG geförderten Projekts² wurden Kompetenzen von Schülern der 6. Klasse im Bereich ‚Wahrnehmen und Kontextualisieren von Musik‘ erforscht. Ausgangspunkt der Überlegun-

2 Das Projekt wurde 2007 bis 2009 an der Universität Bremen durchgeführt. Den Antrag verfassten Andreas Lehmann-Wermser (Universität Bremen), Andreas C. Lehmann (Hochschule für Musik Würzburg) und Anne Niessen (Hochschule für Musik und Tanz Köln), die Mitarbeiter waren Jens Knigge und Anne-Katrin Jordan (beide Universität Bremen). Weitere Informationen zum Projekt sind unter der Adresse <http://www.musik.uni-bremen.de/forschung/komus.html> erhältlich.

gen war die Frage, über welche Kompetenzen Schüler in Bezug auf ihre Hörwahrnehmung von Musik verfügen und wie sie sie aufbauen. Dabei erschien es sinnvoll, von einer Stufung der entsprechenden Kompetenzen auszugehen, weil Wahrnehmungsprozesse sich beispielsweise im Grad der Differenzierung, im Grad der Vernetzung, der Bewusstheit und der Verknüpfung mit Wissensbeständen unterscheiden. Diese Unterschiede beim ‚Wahrnehmen und Kontextualisieren von Musik‘ wurden in einem Kompetenzmodell abgebildet, das auf curricularen Analysen, theoretischen Überlegungen sowie empirischen Forschungsergebnissen basiert (Niessen, Lehmann-Wermser, Knigge & Lehmann 2008; Knigge 2010).

Daran anknüpfend wurden Testaufgaben entwickelt, die zunächst in mehreren Feldtests und nach ihrer Revision im Rahmen einer systematischen Pilotierungsstudie validiert wurden. Die endgültige Auswertung der Pilotierung



Abb. 1: Theoretisches *KoMus*-Kompetenzmodell

steht noch aus;³ sie wird das in Abbildung 1 dargestellte Modell voraussichtlich noch verändern.⁴

Wie funktioniert aber die Messung von Kompetenzen und die Formulierung von Kompetenzniveaus? Gemäß der Weinertschen Definition, die auch im Rahmen von *KoMus* Verwendung fand, sind Kompetenzen „die bei Individuen verfügbaren oder durch sie erlernbaren kognitiven Fähigkeiten und Fertigkeiten, um bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, um die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können“ (Weinert 2001, S. 27). Wenn Problemlösefähigkeiten und -fertigkeiten gemessen werden sollen, muss also das Beobachten von Handeln Rückschlüsse auf die Kompetenzen zulassen und aus diesen Beobachtungen wiederum kann man Testinhalte ableiten (Klieme & Hartig 2007, S. 24).⁵ Kompetenzen sind zudem Dispositionen, die sich nicht auf der Grundlage einzelner Beobachtungen bestimmen lassen. Nötig sind viele Beobachtungsergebnisse bei variierenden Aufgaben und Situationen. „Die konsistente Zusammenfassung solcher Einzelbeobachtungen zu einer Aussage über das individuelle Kompetenzniveau ist das, was in der psychometrischen Fachsprache

-
- 3 Erste Ergebnisse sind bei Jordan, Lehmann & Knigge (2010) dargestellt.
 - 4 Das ursprünglich von Niessen et al. (2008) vorgestellte Modell hat bereits im Zuge der Formulierung des Testkonstrukts (vgl. Abschnitt 2.1) und der daran anschließenden Aufgabenentwicklung verschiedene Modifikationen erfahren. Auf Basis der Ergebnisse der Pilotierungsstudie sind weitere Veränderungen und Präzisierungen, vor allem hinsichtlich der Niveaustuktur zu erwarten.
 - 5 Der Begriff des ‚Problemlösens‘ wird z. T. fälschlicherweise missverstanden als das Lösen von lebensweltlich relevanten Problemen. Der Begriff wird von Weinert jedoch in seiner psychologischen Konnotation verwendet, der sich zunächst nicht auf einen speziellen Kontext bezieht. Verwirrend mag in diesem Zusammenhang auch die Vermischung von Kompetenzbegriff und Literacy-Konzept sein (z. B. in den PISA-Studien), welches sich explizit auf die Lösung von (berufs-)alltagsrelevanten Problemen bezieht. Da ästhetisch geprägte Schulfächer nicht auf im Alltag im engeren Sinne notwendige Anforderungen vorbereiten, erscheint eine Verknüpfung von Kompetenz- und Literacy-Konzept in Bezug auf das Fach Musik nicht sinnvoll. Wir schließen uns daher den Kollegen der Mathematikdidaktik an, die für didaktische Zusammenhänge das Begriffsverständnis von ‚Problemlösen‘ als „Anforderungen bewältigen“ (Büchter & Leuders 2005, S. 188) vorschlagen (vgl. auch die Ausführungen zum Problemlösen im Musikunterricht von Cvetko & Meyer 2009; Niessen 2008).

als Messung bezeichnet wird. Kompetenzmessung hat also nicht, wie es in manchen erziehungswissenschaftlichen Kommentaren immer noch scheint, mit einer ‚Normierung‘ von Gedanken zu tun oder mit einem bloßen Abzählen von Richtigantworten. Messungen können in durchaus komplexen Aussagen resultieren“ (Klieme & Hartig 2007, S. 24).

Ziel des *KoMus*-Projekts war eine möglichst differenzierte Modellierung der Wahrnehmung und Kontextualisierung von Musik. Zur Entwicklung und Validierung eines Kompetenzmodells und eines darauf bezogenen Testinstruments wurde ein Drei-Phasen-Design gewählt: (1) Erstellung eines theoretischen Kompetenzmodells (Niessen et al. 2008), (2) Operationalisierung des Modells in Form von Testaufgaben (Knigge 2010), (3) Empirische Validierung des Modells im Rahmen einer systematischen Pilotierungsstudie (Jordan, Lehmann & Knigge 2010). In vorliegendem Beitrag steht die Phase der Operationalisierung des Modells im Vordergrund. Diese Phase lässt sich in methodischer Hinsicht in zwei Abschnitte unterteilen: Aufgabenentwicklung und Aufgabenanalyse.

2 Theoretische und methodische Grundlagen der Aufgabenentwicklung

2.1 Testkonstrukt: Vom Modell zu den Aufgaben

Im Rahmen des *KoMus*-Projekts wurde zunächst eine Modellskizze entworfen, die auf fachdidaktischem Erfahrungswissen basiert und eine möglichst plausible Dimensionierung und Graduierung der Kompetenz beinhaltet (vgl. Niessen et al. 2008). Um eine Operationalisierung des Modells zu ermöglichen, muss in einem Testkonstrukt möglichst präzise beschrieben werden, was genau unter der Kompetenz ‚Musik wahrnehmen und kontextualisieren‘ verstanden wird und welche Aspekte der Kompetenz durch den Test erfasst bzw. nicht erfasst werden. Das Testkonstrukt dient dabei einerseits einer theoretischen und empirischen Fundierung im Rahmen vorhandener (musikpsychologischer) Forschungen, andererseits können in diesem Zusammenhang die im theoretischen Modell noch relativ abstrakt formulierten Kompetenzdimensionen und -facetten ausgeschärft und konkretisiert werden. Denn erst wenn man dies „elaboriert hat [...], wird man aus der Konstruktdefinition Verhaltensweisen ableiten können, die bei hohen oder geringen Ausprägungen auf dem Konstrukt beobachtbar sein sollten“ (Köller, 2008, S. 166). Die mit einem Modell bzw. Test anvisierte Schülerschaft bestimmt den Rahmen, innerhalb dessen ein

valides Testinstrument entwickelt werden kann. Die Aufgabenentwicklung orientierte sich in *KoMus* an Schülern der sechsten Jahrgangsstufe im Alter von etwa elf bis zwölf Jahren sowie an deren Erfahrungen und musikbezogenen Entwicklungsstand, und sie musste auf die Lernerfahrungen im Fach Musik ausgelegt sein. Aus diesem Grund basiert das Testkonstrukt sowohl auf musikpsychologischen Befunden als auch auf Curriculaanalysen. Das Testkonstrukt bildet somit die Gelenkstelle zwischen dem theoretischen Modell und dessen empirischer Umsetzung in Form von Testaufgaben.

Zusammenfassend ist die Kompetenz des Wahrnehmens und Kontextualisierens von Musik im Testkonstrukt mehrdimensional definiert.⁶ Sie wird als ein Zusammenspiel von Hörwahrnehmungsfähigkeit und dem reflektierten Einsatz musikbezogener Wissensbestände verstanden. Durch die Berücksichtigung musikpsychologischer Befunde (z. B. Gembris 2005; Runfolà & Swanwick 2002) ist einerseits sichergestellt, dass das zu entwickelnde Modell bzw. Testinstrument dem Entwicklungsstand der Schüler angepasst ist. Darüber hinaus kann das Verständnis musikbezogener Wahrnehmung, wie sie im *KoMus*-Projekt modelliert wird, auf Basis musikpsychologischer Grundlagenforschung konkretisiert werden (z. B. Bruhn 2005; Kreutz 2005; La Motte-Haber 2005; Lange 2005; Nauck-Börner 1987; Stoffer 2005). Durch Curriculaanalysen ist das Testkonstrukt in der unterrichtlichen Praxis verankert, wodurch eine curricular-inhaltliche Validität des Modells und der darauf bezogenen Testaufgaben angestrebt wird.

2.2 Testaufgaben: Funktionen, Gütekriterien und Formate

In der pädagogisch-psychologischen und fachdidaktischen Literatur werden verschiedene Aufgabensystematiken vorgeschlagen (z. B. hinsichtlich Inhalt, Funktion oder Format der Aufgaben). Für unseren Zusammenhang ist vor allem die funktionale Unterscheidung von Lern- und Testaufgaben wichtig (z. B. Benner 2007; Caspari, Grotjahn & Kleppin 2008). Während bei Lernaufgaben der Anregungsgehalt und das Lernpotenzial im Vordergrund stehen, besteht die Funktion von Testaufgaben darin, Kompetenzen (oder andere Merkmale) einer empirischen Überprüfung zugänglich zu machen. In diesem Sinne werden Testaufgaben zur Leistungsüberprüfung in Schulleistungstudien, nicht zuletzt aber auch in Klassen- und Abschlussarbeiten eingesetzt. Dieser Praxis liegt die Annahme zugrunde, dass aus dem Lösen von Aufgaben

6 Detaillierte Ausführungen zum Testkonstrukt finden sich bei Knigge (2010) und Jordan et al.(i. Vorb.).

mit einer relativ hohen Sicherheit auf das Vorhandensein bzw. Fehlen der entsprechenden Kompetenzen bei Schülern geschlossen werden kann. Um das leisten zu können, müssen Testaufgaben bestimmten formalen und psychometrischen Kriterien genügen.

Die im *KoMus*-Projekt entwickelten Testaufgaben erfüllen zwei Funktionen: (1) Mittels der Aufgaben soll eine empirische Überprüfung und ggf. notwendige Modifikation des Kompetenzmodells, auf das sich die Aufgaben beziehen, erfolgen; (2) durch Analyseverfahren und Itemselektion (s. Abschnitt 3) soll aus den Aufgaben ein standardisiertes Testinstrument gebildet werden. Hierfür ist es erforderlich, dass die Aufgaben bestimmten Qualitätsansprüchen genügen. Neben den Hauptgütekriterien von Tests (Objektivität, Reliabilität und Validität; s. z. B. Moosbrugger & Kelava 2007) ist auch die Wahl eines geeigneten Aufgabenformats entscheidend für die optimale Erfassung einer anvisierten Kompetenz. Für die Entwicklung von Testaufgaben stehen verschiedene Item-Typen und -Formate zur Verfügung (vgl. auch Knigge 2010, Kap. 3.1). Jedes der Formate hat Vor- und Nachteile. Geschlossene Formate sind sehr ökonomisch in der Bearbeitung und Auswertung bei gleichzeitig maximaler Auswertungsobjektivität. Komplexe und kreative Fähigkeiten können aber oft nur schwer oder gar nicht mit geschlossenen Formaten erfasst werden (Rost 2004, S. 59 ff.). Offene Aufgaben sind hingegen eher geeignet für komplexere Anforderungen und das Antwortverhalten lässt sich leichter auf reale Situationen übertragen. ‚Erkauft‘ wird dies jedoch mit einem relativ hohen Zeitaufwand bei der Bearbeitung der Aufgaben und vor allem bei der Auswertung. Es muss daher darauf geachtet werden, dass das gewählte Aufgabenformat und die anvisierte Kompetenz in einem entsprechenden Passungsverhältnis zueinander stehen, um eine ökonomische, aber gleichzeitig möglichst objektive und valide Messung zu gewährleisten.

3 Prozess der Aufgabenentwicklung im Projekt *KoMus*⁷

Auf Basis des Testkonstrukts konnte die Operationalisierung des Modells erfolgen. Ein Entwicklungsteam, bestehend aus sechs Kooperationslehrern und den Wissenschaftlern des *KoMus*-Projekts führte von Februar bis Dezember 2008 zehn Sitzungen zur Entwicklung von Testaufgaben durch. Der Prozess war zirkulär konzipiert: (1) monatliche Sitzung des Entwicklungsteams, (2)

7 Ausführlicher ist der Prozess der Aufgabenentwicklung bei Knigge 2010 (v. a. Kap. 4.3) dargestellt.

Erprobung (Feldtest) der Aufgaben in den Klassen der Kooperationslehrer, (3) Auswertung des Tests, (4) nächste Entwicklungssitzung: Überarbeitung der Aufgaben und Entwicklung neuer Aufgaben. Um eine möglichst effiziente und an den Testgütekriterien orientierte Aufgabenentwicklung sicherzustellen, wurde ein Handbuch mit ausführlichen Hinweisen zur Aufgabenkonstruktion erstellt (Knigge 2008).⁸ Darin waren u. a. folgende Vorgaben festgelegt:

- Handlungsleitend für die jeweilige Aufgabenentwicklung sollte die vorab zu treffende Entscheidung sein, welche (Teil-)Kompetenz mit einer Aufgabe erfasst werden soll. Jeder Aufgabenentwurf war dementsprechend mit einer Beschreibung der intendierten Kompetenzmessung (Dimension und Niveau) zu versehen.
- Es sollten ca. 25% offene, 25% halb-offene und 50% geschlossene Items verwendet werden, wobei halb-offene und insbesondere offene Items hauptsächlich für komplexere Anforderungen auf höheren Kompetenzniveaus vorgesehen waren.
- Hörbeispiele sollten ein breites stilistisches Spektrum abdecken und möglichst nicht länger als 20-30 sec. sein.
- Für jede Aufgabe waren die Lösungen zu dokumentieren (wichtig v. a. bei offenen Formaten).

Auf dieser Basis sollten insgesamt mind. 120 Items entwickelt und überprüft werden.⁹ Um im Rahmen der Tests eine möglichst objektive Testdurchführung zu gewährleisten, wurden die Aufgabenentwürfe (Ø ca. 28 Items pro Sitzung) im Anschluss an die Entwicklungssitzungen in ein standardisiertes Testheft übertragen und zusammen mit einer Audio-CD und Anweisungen für die Testdurchführung (Testleiter-Manual) an die Kooperationslehrer zur Erprobung in ihren Klassen versandt. Zusätzlich erhielten Schüler und Lehrer einen Rückmeldebogen, sodass Verständnisschwierigkeiten und sonstige Probleme bei der Testdurchführung sofort festgehalten werden konnten.

8 Das Handbuch ist eine weiterentwickelte und speziell auf KoMus abgestimmte Fassung von Köller et al. 2005.

9 Zugrunde gelegt wurde hierbei die hypothetische Struktur des Modells von drei Kompetenzniveaus je Modelldimension (vgl. Niessen et al. 2008). Die avisierte Itemanzahl ergibt sich somit aus zehn Items pro Dimension und Niveau. So eine relativ große Anzahl von Items ist im Speziellen bei einer Neuentwicklung eines Testinstruments vonnöten, da im Zuge der Itemselektion ein gewisser Teil der Items aufgrund statistischer Kriterien verworfen werden muss.

4 Aufgabenanalyse

4.1 Itemselektion

Dank einer relativ hohen Schülerbeteiligung (\varnothing pro Test $N = 215$) konnten bereits im Rahmen der Aufgabenentwicklung umfangreiche statistische Analysen durchgeführt werden. Mittels einer Kombination aus klassischen und probabilistischen¹⁰ statistischen Verfahren, aber auch mithilfe der Anwendung von qualitativen Methoden wurden die entwickelten Aufgaben in *KoMus* systematisch auf ihre Messeigenschaften hin untersucht. Zunächst ging es hierbei im Sinne der Itemselektion darum, die besten Items für die Modellvalidierung bzw. für das zu erstellende Testinstrument zu identifizieren. Verschiedene Analysemethoden stellten die Grundlage dar, auf der Items ausgewählt, überarbeitet oder verworfen wurden. Tabelle 1 fasst die wichtigsten Analysen und die dabei angewandten psychometrischen Kriterien zusammen.¹¹

Analyse	Kriterium
(Klassische) Itemschwierigkeit	$95 > P_i > 5$
Distraktoren	Alle Distraktoren ungefähr gleich häufig gewählt; Trennschärfe der Distraktoren $< .05$
Trennschärfe	$r_{it} \geq .25$
Itemfit (Rasch-Modell)	$1.20 > \text{MNSQ} > 0.80$
Differential Item Functioning (DIF)	Für ein Item liegt kein substantielles DIF vor ¹²

Tabelle 1 Durchgeführte Itemanalysen im Rahmen der Aufgabenentwicklung

10 Zur Probabilistischen Testtheorie vgl. z. B. Rost (2004) und Bühner (2006). Die im Rahmen der Aufgabenentwicklung und -analyse eingesetzten probabilistischen Verfahren und Modelle (insbesondere das Rasch-Modell) sind bei Knigge (2010, Kap. 5) beschrieben.

11 Für eine ausführlichere Darstellung aller Analysemethoden und Kriterien s. Knigge 2010, Kap. 5 u. 6.1.

12 Als substantielles ‚Differential Item Functioning‘ wurden signifikante DIF-Werte > 0.50 logits definiert (vgl. Wang 2000).

Items mit zufriedenstellenden statistischen Werten und unproblematischer Bearbeitung wurden in eine Item-Datenbank übernommen („selektiert“). Die übrigen Items wurden erneut im Rahmen der nächsten Entwicklungssitzung diskutiert und ggf. überarbeitet oder verworfen. Durch diesen zirkulären Prozess der Aufgabenerstellung, -erprobung und -überarbeitung konnte ein psychometrisch hochwertiger Itempool (179 Items) generiert werden, der die Strukturen des theoretischen Kompetenzmodells abbildet, auf unterrichtlicher und curricularer Ebene verankert ist und dabei sowohl eine Differenzierung über das gesamte Fähigkeitsspektrum (von Schülern der sechsten Jahrgangsstufe) ermöglicht als auch den strengen testtheoretischen Annahmen des Rasch-Modells genügt (vgl. Knigge 2010, Kap. 6.1). Die zentralen psychometrischen Kennwerte des Itempools sind in Tabelle 2 im Anhang dargestellt.

Des Weiteren war es ein zentrales Anliegen des Forschungsprojekts, über die statistischen Kennwerte hinaus mehr über die Messeigenschaften der entwickelten Aufgaben zu erfahren, da so ein besseres Verständnis der Strukturen der zu erfassenden Kompetenz erlangt werden kann. Konkret handelt es sich hierbei um die Identifikation sogenannter „schwierigkeitsgenerierender Aufgabenmerkmale“, die im Folgenden ausführlicher dargestellt werden.

4.2 Identifikation von schwierigkeitsgenerierenden Aufgabenmerkmalen¹³

Als „schwierigkeitsgenerierende Aufgabenmerkmale“ werden die Eigenschaften einer Testaufgabe bezeichnet, „die mit höheren oder niedrigeren Anforderungen an die getesteten Personen einhergehen und damit die Lösungswahrscheinlichkeiten der Aufgaben beeinflussen“ (Hartig & Jude 2007, S. 31). Bei der Identifikation solcher Aufgabenmerkmale geht es folglich darum, die Charakteristika einer Aufgabe zu beschreiben, die in Bezug auf die Itemschwierigkeit relevant sind. Im Rahmen von *KoMus* wurde der Versuch unternommen, schwierigkeitsgenerierende Aufgabenmerkmale zu identifizieren, die sich auf die *Aufgabe*, das *Hörbeispiel*, den *Notentext* (sofern vorhanden), die *Wahrnehmungsanforderungen* und die für die Aufgabenlösung notwendige *Wissensbasis* beziehen. An dieser Stelle sollen zusammenfassend Ergebnisse dieser Analysen vorgestellt werden.

13 Die folgenden Abschnitte sind eine Zusammenfassung der bei Knigge (2010, Kap. 7) durchgeführten Analysen.

4.2.1 Vertiefende Analysen der Items zur Rhythmuswahrnehmung

In den meisten Curricula spielt die Rhythmuswahrnehmung eine wichtige Rolle. Schüler sollen sich u. a. Rhythmen merken können, diese wiedererkennen, aus motivischem Material Rhythmen extrahieren oder Rhythmen einem Notenbild zuordnen. Im Testkonstrukt wurde dementsprechend die Rhythmuswahrnehmung als eine Facette der Wahrnehmungskompetenz definiert und anschließend operationalisiert.

Die durchgeführten Analysen zeigen, dass eine vertiefende inhaltliche Betrachtung der Rhythmus-Items und deren empirisch ermittelter Schwierigkeit sehr aufschlussreich ist in Bezug auf die Aufgabencharakteristika und die darauf bezogenen Lösungsprozesse. Auf Basis der Analysen können die identifizierten schwierigkeitsgenerierenden Merkmale zusammenfassend folgendermaßen beschrieben werden (Knigge 2010, Kap. 7.1):

Die Schwierigkeit eines Rhythmus-Items ist abhängig von

- der Komplexität der klanglichen Struktur, in der ein Rhythmus identifiziert werden muss,
- der Komplexität der rhythmischen Struktur,
- dem notwendigen Wissen in Bezug auf Notation,
- der Anwendungsform von Notationskenntnissen,
- den Anforderungen an das musikalische Gedächtnis.

Zusätzlich können auch Merkmale identifiziert werden, die die Aufgabenlösung erleichtern: Wenn die klangliche oder melodische Struktur die rhythmische Struktur unterstützen und somit einprägsamer und leichter memorierbar machen, wirkt dies der Itemschwierigkeit entgegen.

4.2.2 Vertiefende Analysen der Items zur Formwahrnehmung

Eine weitere Facette des Kompetenzmodells bezieht sich auf die Wahrnehmung von musikalischen Formverläufen. Die Formwahrnehmung gehört, ebenso wie die Rhythmuswahrnehmung, zu den zentralen Fähigkeiten, deren Aufbau auf curricularer Ebene gefordert wird. Dort lässt sich auch ein breiter Konsens in Bezug auf die konkreten Inhalte ausmachen. Häufig wird dabei unterschieden zwischen elementaren Form-/Gestaltungsprinzipien (z. B. Wiederholung, Variation) und Formmodellen (z. B. Rondo). Für die Aufgabenentwicklung wurden nur Formen verwendet, die durch einen Großteil der Curri-

cula abgedeckt sind. Um darüber hinaus die unterrichtliche Relevanz der gewählten Formen noch weiter abzusichern, konnten die Ergebnisse von Schulbuchanalysen verwendet werden. Unter Berücksichtigung beider Quellen erfolgte die Auswahl von ‚Kanon‘, ‚Rondo‘ und ‚Liedformen‘. Insgesamt wurden 13 Items zur Formwahrnehmung entwickelt, die sich größtenteils explizit auf ein bestimmtes Formmodell beziehen, teilweise aber auch die Wahrnehmung einzelner Formprinzipien fokussieren.

Mittels verschiedener Analysen können auch in Bezug auf die Items zur Formwahrnehmung eine Reihe von schwierigkeitsgenerierenden Aufgabenmerkmalen identifiziert werden (Knigge 2010, Kap. 7.2). Es lassen sich vier Merkmale formulieren, für die von einem Einfluss auf die Itemschwierigkeit auszugehen ist. Demgemäß ist die Schwierigkeit der Items zur Erfassung von Formwahrnehmung abhängig von

- der physischen Markierung von Abschnitten. Es wird angenommen, dass die Itemschwierigkeit steigt, umso weniger deutlich physische Markierungen die Segmentgrenzen zwischen zwei Abschnitten kennzeichnen;
- der notwendigen Nutzung des musikalischen Gedächtnisses. Sofern ein Formabschnitt (oder Teile davon) memoriert werden müssen, ist von einer erhöhten Itemschwierigkeit auszugehen;¹⁴
- der Länge und Komplexität der Formteile;¹⁵
- dem für die Aufgabenlösung notwendigen expliziten Wissen.

14 In Bezug auf den KoMus-Itempool kennzeichnet dieses Merkmal den Unterschied zwischen der Erkennung von Abschnitten und dem Vergleich von Abschnitten. Dies gilt selbstverständlich nur für diese Items und die Art, wie dort Formwahrnehmung erfasst wird. Da bei den KoMus-Items zur Abschnitterkennung die Segmentgrenzen immer durch relativ deutliche physische Markierungen gekennzeichnet sind, spielt hierbei das musikalische Gedächtnis keine oder nur eine untergeordnete Rolle. Bei komplexeren Formen (z. B. Fuge, Variation) ist es hingegen häufig nicht möglich die Segmentgrenzen zu bestimmen, ohne z. B. ein Thema, Motiv o. ä. zu memorieren.

15 Ein Einfluss dieses Merkmals wird nur für Aufgaben vermutet, die einen Vergleich von Formteilen verlangen. Sofern dieser nicht notwendig ist und auch zur Bestimmung der Segmentgrenzen kein musikalisches Gedächtnis erforderlich ist, sollte das Merkmal nicht schwierighkeitsrelevant sein.

In diesem und dem vorangegangenen Abschnitt konnten einerseits verschiedene Merkmale beschrieben werden, die in genuinem Zusammenhang mit der Wahrnehmung von Rhythmus und Form stehen. Andererseits erbrachten die Analysen aber auch Merkmale, die sich auf grundlegende Wahrnehmungsvorgänge, Eigenschaften eines Hörbeispiels oder die Art einer Aufgabenstellung beziehen, also vermutlich nicht ausschließlich auf die Rhythmus- und Formwahrnehmung beschränkt sind, sondern auf übergeordneter Ebene anzusiedeln sind. Auf diese Gruppe von schwierigkeitsgenerierenden Merkmalen soll in den folgenden beiden Abschnitten näher eingegangen werden.

4.2.3 Wissensbasierte Aufgabenmerkmale

Zunächst mag es verwundern, dass explizites Fachwissen im Rahmen eines Kompetenzmodells bzw. darauf bezogener Testaufgaben eine größere Rolle spielt. Wird doch in der Diskussion um Input- und Outputsteuerung immer wieder betont, dass mit dem Kompetenzbegriff eine Bewegung weg von der Vermittlung einzelner Inhalte, von der Fokussierung auf Faktenwissen, hin zu einem an Fähigkeiten und Fertigkeiten orientieren Lehr-/Lernkonzept verbunden ist. Betrachtet man jedoch den Kompetenzbegriff genauer (vgl. Abschnitt 1), so wird deutlich, dass dieser keineswegs den Stellenwert von Wissen grundsätzlich infrage stellt. Vielmehr geht es darum, den Erwerb von Wissen in einen größeren und vor allem anwendungsbezogenen Zusammenhang zu stellen: „Kompetenz stellt die Verbindung zwischen Wissen und Können [...] her und ist als Befähigung zur Bewältigung von Situationen bzw. von Aufgaben zu sehen“ (Klieme et al., 2003, S. 73). Wissen ist also eine zentrale Facette von Kompetenz, die jedoch keinen Wert an sich hat, sondern eher im Sinne einer Ressource verstanden wird, die Schüler in die Lage versetzt, mit den an sie gestellten Handlungsanforderungen sinnvoll umgehen zu können (Criblez et al. 2009, S. 36).

Sofern ein Kompetenzmodell und ein darauf bezogenes Testverfahren beanspruchen auf der theoretischen Basis des Kompetenzbegriffs konstruiert und curricular valide zu sein, so ist es sinnvoll und notwendig auch explizites Fachwissen bei der Konstruktion von Testaufgaben mit einzubeziehen. Im Anschluss an die vorigen Ausführungen ist hierfür entscheidend, dass es sich dabei nicht um die isolierte Abfrage einzelner Inhalte handelt (z. B. „Wann wurde Mozart geboren?“), sondern das Wissen in konkreten Anforderungssituationen angewandt werden muss. Dieser Grundsatz wurde bei der Konstruktion der *KoMus*-Items eingehalten, sodass der Einsatz von Wissen bei einem Item immer nur ein Aufgabenmerkmal unter anderen ist.

Die durchgeführten Analysen zu wissensbasierten Aufgabenmerkmalen zeigen, dass grundsätzlich von einem schwierigkeitsrelevanten Einfluss auszugehen ist, wenn für die Lösung einer Aufgabe der Einsatz von Fachwissen notwendig ist. Das Merkmal ‚Einsatz von Fachwissen‘ konnte außerdem noch weiter ausdifferenziert werden (Knigge 2010, Kap. 7.3):

- Die Itemschwierigkeit ist nicht nur davon abhängig, ob das Merkmal grundsätzlich vorliegt, auch die Qualität des Wissens ist von Relevanz. Demgemäß ist von einer steigenden Itemschwierigkeit auszugehen, umso detaillierter und elaborierter das Wissen vorhanden sein muss.
- In inhaltlicher Hinsicht erscheint die Aufschlüsselung des Merkmals nach Wissensdimensionen sinnvoll. Es ergeben sich somit vier wissensbasierte Merkmale, die sich in Bezug auf die inhaltliche Dimension unterscheiden (musiktheoretisches, -historisches und -stilistisches Fachwissen sowie Wissen in Bezug auf kulturelle und soziale Kontexte von Musik).

4.2.4 Merkmalsebene ‚Aufgabe‘

In Studien, die den Einfluss von Aufgabenmerkmalen auf die Itemschwierigkeit untersuchen, werden häufig die verwendeten Itemformate als Merkmale beschrieben (z. B. Prenzel et al. 2002). Hinter dieser Vorgehensweise steht die Annahme, dass die Schwierigkeit, beispielsweise einer Mathematik-Aufgabe, nicht nur von den mathematischen Kompetenzen der Schüler abhängt, sondern auch durch das Itemformat beeinflusst wird. Es wird angenommen, dass es Schülern grundsätzlich schwerer fällt, eine Antwort eigenständig zu formulieren (freie Formate: halboffen, offen), als eine vorgegebene Antwortalternative auszuwählen (geschlossene Formate: Multiple-Choice-, Richtig-Falsch-, Zuordnungs-Items). Hierfür sind in der Regel zwei Faktoren ausschlaggebend. Einerseits stellt ein freies Format erhöhte Anforderungen an die sprachlichen Fähigkeiten (Textproduktion/Schreibleistung). Andererseits spielen auch motivationale Aspekte eine Rolle, denn ein freies Aufgabenformat ist in Bezug auf die rein technische Bearbeitung immer aufwendiger zu lösen als ein geschlossenes Format, bei dem z. B. lediglich eine Antwort angekreuzt werden muss. Ein weiteres Merkmal, das auf der Ebene der technischen Oberflächencharakteristika einer Aufgabe anzusiedeln ist, sind die sprachlichen Anforderungen, die durch die textspezifische Beschaffenheit des Itemstamms, aber auch der Antwortalternativen gegeben sind.

Auf Basis verschiedener Itemanalysen werden auf der Merkmalsebene ‚Aufgabe‘ folgende Merkmale spezifiziert (Knigge 2010, Kap. 7.4):

- Das Merkmal ‚Itemformat‘ umfasst drei Ausprägungen: geschlossene, halb-offene und offene Formate.
- Das Merkmal ‚Textlänge‘ bezieht sich sowohl auf den Itemstamm als auch auf die Antwortalternativen. Die Items werden dabei klassifiziert in Items mit viel und wenig Text. In Bezug auf die KoMus-Items genügt solch eine dichotome Kodierung, da sich die Items in der Regel sehr deutlich hinsichtlich des Merkmals unterscheiden.
- Das Merkmal ‚formalsprachliche Anforderungen‘ bezieht sich auf das in einer Aufgabe verwendete Vokabular (hochfrequente Wörter, weniger frequente Wörter, erweiterter Wortschatz) und die grammatikalischen Strukturen (einfache syntaktische Strukturen, komplexere Strukturen). Die Ausprägungen des Merkmals wurden dabei in Anlehnung an Nold & Rossa (2007) formuliert.

Eine Untersuchung des Zusammenhangs von Itemschwierigkeit und der auf Aufgabenebene identifizierten Merkmale ist aufschlussreich und wichtig, da so der Einfluss von diesen allgemeinen und eher technischen Oberflächencharakteristika einer Aufgabe unterschieden werden kann von den Aufgabenmerkmalen, die in genuinem Zusammenhang mit den anvisierten musikspezifischen Kompetenzen stehen. Die Analyse von nicht kompetenzspezifischen Merkmalen dient somit auch der Absicherung der Konstruktvalidität der Items bzw. des Kompetenztests.

4.2.5 Empirische Analyse der Zusammenhänge von Aufgabenmerkmalen und Itemschwierigkeiten

Die durch inhaltliche und vergleichende Analysen identifizierten Aufgabenmerkmale wurden abschließend einer statistischen Überprüfung unterzogen (Knigge 2010, Kap. 7.6). Ein geeignetes Verfahren, um die empirisch ermittelten Itemschwierigkeiten mit den Aufgabenmerkmalen in Beziehung zu setzen, ist die multiple lineare Regressionsanalyse (vgl. Hartig 2007). Hierbei wird untersucht, ob die Unterschiede der Itemschwierigkeiten unter Verwendung der Aufgabenmerkmale erklärt werden können. Das Ausmaß erklärter Unterschiede ist ein Indikator dafür, ob sich die angenommenen schwierigkeitsgenerierenden Merkmale durch die tatsächlichen Aufgabenschwierigkeiten bestätigen lassen. Darüber hinaus kann auf Basis der Regressionsanalysen auch beurteilt werden, ob einzelne Merkmale besonders bedeutsam für die Aufgabenschwierigkeit sind oder aber Merkmale eine eher geringe Erklärungskraft für die Schwierigkeit der Aufgaben besitzen.

Auf Basis der Regressionsanalysen können folgende Befunde berichtet werden:

1. Die Analysen zeigen, dass sich empirisch ein schwierigkeitsgenerierender Effekt der Aufgabenmerkmale auf die Itemschwierigkeiten nachweisen lässt. Die in den Regressionen durchweg hohen Varianzaufklärungen ($.628 \leq R^2_{\text{kor}} \leq .914$) sind aufgrund der Datenbasis mit der gebotenen Vorsicht zu interpretieren, deuten jedoch auf eine starke Vorhersagekraft der Merkmale hin.
2. Der Großteil der identifizierten Merkmale konnte in die Regressionsmodelle einbezogen und dadurch validiert werden. Lediglich die Überprüfung des Einflusses der beiden Merkmale ‚formalsprachliche Anforderungen‘ und ‚Fachwissen – soziale/kulturelle Kontexte‘ war aufgrund der Datenlage nicht möglich und muss weiterführenden Untersuchungen vorbehalten bleiben.
3. Betrachtet man den Einfluss der einzelnen Merkmale, so ist es aufgrund methodischer Grenzen nur bedingt möglich, diese hinsichtlich ihrer Stärke miteinander zu vergleichen. Tendenziell scheint sich aber anzudeuten, dass die Schwierigkeit eines Items primär durch die Anforderungen an die Hörwahrnehmung und das notwendige Fachwissen beeinflusst wird. Ebenso zeigt sich ein schwierigkeitsgenerierender Einfluss der Anforderungen an das musikalische Gedächtnis, der vermutlich aber weniger stark ist.

4.2.6 Fazit

Die Identifikation, detaillierte Beschreibung und empirische Überprüfung von schwierigkeitsgenerierenden Aufgabenmerkmalen ermöglicht ein tieferes Verständnis der zur Aufgabebearbeitung notwendigen Prozesse und damit des anvisierten Kompetenzkonstrukts (vgl. auch Hartig & Jude, 2007, S. 31). Dieses Verständnis erlaubt eine präzisere Interpretation von Testdaten sowie eine differenzierte Formulierung eines Kompetenzmodells und kann nicht zuletzt für die zukünftige Konstruktion von Items und Tests genutzt werden.

5 Ausblick: Der Einsatz von ‚Cognitive Labs‘

Abschließend soll eine Methode skizziert werden, die die Entwicklung und Auswertung von Testaufgaben unterstützen kann. Bei der Arbeit des *KoMus*-Aufgabenentwicklungsteams wurde immer wieder deutlich, wie wenig eigentlich darüber bekannt ist, auf welche Weise Schülerinnen und Schüler Aufga-

ben lösen. Die Strategien der Schüler zu kennen, wäre aber sowohl für die Konstruktion der Aufgaben wie für die Bestimmung der jeweiligen Kompetenzen wichtig gewesen. Ein qualitatives Verfahren für die Untersuchung von Lösungsstrategien stellt die ‚Cognitive-Lab’-Methode dar (z. B. Ericsson & Simon 1999). ‚Cognitive Labs’ waren nicht Bestandteil der systematischen Aufgabenentwicklung im Rahmen von *KoMus*, vielmehr wurde diese Methode nur explorativ in Bezug auf einzelne Items angewendet. Im Ausblick am Ende dieses Beitrags möchten wir an einer Beispielaufgabe demonstrieren, wie auf diese Weise weitere Informationen über die Kompetenzen von Schülern und die Messeigenschaften von Items gewonnen werden können.

5.1 Begründung der Aufgabenwahl

Bei einer der *KoMus*-Aufgaben wurde den Schülerinnen und Schülern ein eintaktiger Rhythmus mit zwei verschiedenen Notenwerten vorgestellt, den sie anschließend unter insgesamt vier Rhythmen wieder heraushören sollten: eine der Aufgaben, die in Bezug auf Lebensnähe und Komplexität nicht zu den gelungensten gehört – auch wenn man dieses Format gelegentlich in interaktiven Kinderspielen und in den Medien vorfindet. Gerade an diesem einfachen Beispiel lässt sich aber gut demonstrieren, welche Fragen sich bei der Testkonstruktion stellen können. Bei der Erstellung dieses Items wurde im Kreis der Fachdidaktiker und Lehrenden deutlich, dass über die Strategien, die Schüler zu dessen Lösung anwenden würden, nur spekuliert werden konnte: So war nicht klar, ob die Probanden sich den Rhythmus eher beispielsweise mit Hilfe von Abzählen zu merken versuchen oder ob sie den auditiven Eindruck memorieren würden. Die gewählte Lösungsstrategie ist aber bedeutsam für die Aufgabenkonstruktion: Wenn die Probanden den Rhythmus im Arbeitsgedächtnis speichern, wird die Reihenfolge der Items wichtig: Dann nimmt die Schwierigkeit deutlich zu, je nachdem ob die richtige Antwort an erster, zweiter, dritter oder gar vierter Stelle erscheint. Wenn die Probanden jedoch die Schläge zählen oder sich die Struktur verbalisiert merken (z. B. „erst zwei kurze Schläge, dann noch drei lange“), spielt die Reihenfolge der Antwortalternativen keine so wichtige Rolle. Die Aufgabe sollte auf dem mittleren Schwierigkeitsniveau angesiedelt werden; es wurde entschieden, die richtige Antwort an die dritte Stelle zu setzen und die Aufgabe im Rahmen eines ‚Cognitive Labs’ explorativ zu untersuchen.

5.2 *Die Methode der ‚Cognitive Labs‘*

Lautes Denken als übergeordneter Begriff bezeichnet eine empirische Forschungsmethode, bei der Personen aufgefordert werden, „ihre Gedanken laut auszusprechen, während sie sich einer Aufgabe oder Tätigkeit widmen“ (Bilandzic 2005, S. 362). Die Methode stammt als ‚thinking aloud technique‘ aus der US-amerikanischen Psychologie (u. a. Ericsson & Simon 1984). Aus den Audio- und Videoaufnahmen solcher Erhebungssituationen werden ‚verbal protocols‘ erstellt, die dann ausgewertet werden können. Ericsson und Simon unterscheiden grundsätzlich zwischen zwei verschiedenen Ergebnisarten des lauten Denkens: Zunächst kann in ‚verbal protocols‘ Denken beschrieben sein, das ohnehin in Form von Sprache vorliegt, ‚vocalisation‘ oder ‚talk aloud‘. Davon unterscheiden sie ‚verbalisation‘ oder ‚think aloud‘, nämlich Denken, das in anderer Form, z. B. als Visualisierung, vorliegt und das erst in Wörter ‚übersetzt‘ werden muss (Nielsen 2002, S. 105).

In den 1980er Jahren entwickelten Forscher im Zusammenhang mit Testverfahren auf der Grundlage der ‚thinking aloud technique‘ die spezifischere Form der ‚cognitive labs‘, weniger häufig auch ‚cognitive interviews‘ genannt. Sie ist charakterisiert durch den Gegenstandsbezug und daraus folgend durch bestimmte methodische und methodologische Besonderheiten: „A cognitive lab is a method of studying the mental processes one uses when completing a task such as solving a mathematics problem or interpreting a passage of text“ (Zucker, Sassmann & Case 2004, S. 2). Besonders geeignet sind ‚Cognitive Labs‘, um die Gründe für Ungereimtheiten und Probleme bei Testungen aufzudecken. So ist es mit ihrer Hilfe möglich, Items, die unterschiedlich interpretiert werden können, zu identifizieren, zu komplizierte oder zu schwierige Anweisungen auszumachen sowie Gründe für alle Arten von erstaunlichen Ergebnissen zu finden (Zucker, Sassmann & Case 2004, S. 4).

5.3 *Der Einsatz der Methode in KoMus*

Für die ‚Cognitive Labs‘, die im Rahmen von *KoMus* in explorativer Absicht durchgeführt wurden, wurde ein Design gewählt, das sich in der Datenerhebung eng an das Verfahren von Zucker et al. (2004) anlehnt:

„In a cognitive lab, a student completes test items and verbally reports his or her thoughts related to the item using a combination of ‘think-aloud’ sessions (concurrent verbal reporting) and interviews with the researcher after each item is completed (retrospec-

tive verbal reporting) ... To obtain a more thorough understanding of the subject's mental processes, verbal reports are often combined with other behavioral data observed during the cognitive lab, such as how a subject uses a pencil and paper to solve a mathematics item“ (S. 4).

Tatsächlich erwies es sich als wichtiges Element der Untersuchung, die Schülerinnen und Schüler genau zu beobachten. So steckte in einer kleinen Verzögerung der Antwort eine wichtige Aussage und so war es möglich, zunächst unverständlich erscheinendes Vorgehen durch Nachfragen zu erhellen.

Die Auswertung der Daten wurde in entscheidenden Punkten anders vorgenommen, als das von Ericsson und Simon (1984) vorgeschlagen wird. Für die Auswertung wurde ein Vorgehen gewählt, das sich an das der Grounded Theory Methodologie (Strauss 1994) anlehnt, in der Annahme, dass es für jeden Aufgabentyp und bei jedem einzelnen Probanden eine je besondere Art und Weise der Lösung geben könnte. Dieses Vorgehen erwies sich z. B. als sinnvoll in Fällen, in denen Probanden von Lösungsmöglichkeiten berichteten, die sie tatsächlich aber gar nicht angewendet hatten. Zudem wurde, anders, als Simon und Ericsson es vorschlagen, in den Auswertungsprozess Kontextwissen einbezogen und die Umstände der Erhebungssituation reflektiert: Die Frage, ob die Probanden über musikpraktische Erfahrung verfügten, erschien ebenso wichtig wie der Einbezug der Qualität des Ergebnisses. Und: Die Analyse erfolgte nicht anhand eines vorgefertigten Kategoriensystems, wie es von Ericsson und Simon empfohlen wird. Tatsächlich sollten ja gerade unerwartete Lösungsstrategien identifiziert und den Äußerungen der Probanden möglichst unvoreingenommen begegnet werden.

Es wurden ca. 20 Schülerinnen und Schüler im Alter von 10 bis 13 Jahren in ‚Cognitive Labs‘ befragt, allerdings liegen die Ergebnisse in unterschiedlich gut dokumentierter Form vor: ‚Cognitive Labs‘ mit etwa 15 Schülern fanden im Rahmen eines Universitätsseminars statt und wurden nur mit Hilfe von Gesprächsprotokollen festgehalten; zudem wurden dort jeweils drei bis vier Probanden in kleineren Gruppen befragt, so dass von einer gegenseitigen Beeinflussung durch die Äußerungen der Mitschüler ausgegangen werden muss. Aus den ‚Cognitive Labs‘ mit fünf weiteren Probanden liegen wörtliche Transkriptionen vor, wobei zwei davon sehr ausführlich sind. In diesen beiden letzten ‚Cognitive Labs‘ wurden die Lösungen von nur zwei Items begleitet, was ausführlichere Gespräche über die jeweiligen Items erlaubte. Die Auswertung der Daten erfolgte schließlich mit Hilfe eines mehrstufigen Kodierverfah-

rens, das dem Vorgehen im Rahmen der Grounded Theory ähnelt, ohne dass insgesamt im Sinne dieses Ansatzes geforscht worden wäre: Im ersten Schritt wurde kleinschrittig und möglichst ‚offen‘ kodiert, im zweiten Schritt wurden die Kodierungen systematisiert, verglichen und teilweise stärker miteinander in Beziehung gesetzt. Im dritten Arbeitsschritt wurden dann die Lösungswege der Schüler und Schülerinnen rekonstruiert. Dabei wurde berücksichtigt, ob die Probanden bei der Aufgabenlösung erfolgreich waren und ob sie über längere musikpraktische Erfahrung¹⁶ verfügten.

Zunächst schien die Suche nach Strategien zur Lösung der Rhythmusaufgabe eine beeindruckende Fülle von Möglichkeiten zu Tage zu fördern. Bei genauerem Hinsehen reduziert sich aber die Vielfalt und lässt sich in folgenden Thesen kondensieren:

- Es gibt nur eine Strategie, die relativ zuverlässig die Lösung der Aufgabe ermöglicht: Sie besteht aus dem Memorieren des Rhythmus in einer nicht-verbalisierten Form. Von dieser Strategie berichten im Rahmen der ‚Cognitive Labs‘ vor allem Schüler, die über musikpraktische Vorbildung verfügen.
- Die kognitiven Strategien, die außerdem beschrieben werden, also das verbalisierte Memorieren von Folgen längerer oder kürzerer Notenwerte, werden eher von Schülern ohne intensive musikpraktische Erfahrungen beschrieben und führen bei ihnen nicht zum Erfolg.

5.4 Ausblick

Für die Interpretation der Ergebnisse der ‚Cognitive Labs‘ ist es sinnvoll, die Daten der *KoMus*-Pilotierungsstudie ergänzend hinzuzuziehen: Insgesamt 59,8 % der Sechstklässler (N = 508) lösten die Aufgabe erfolgreich. Schüler, die mindestens 2 Jahre Instrumentalunterricht hatten, antworteten signifikant häufiger richtig (65,6 %) als Schüler, die keinen oder weniger als ein Jahr Unterricht hatten (53,5 %). Allerdings kann von einer Signifikanz nur auf der Basis

16 Auch wenn diese Etikettierung nicht unproblematisch ist: Unter „längerer musikpraktischer Erfahrung“ wird im Folgenden verstanden, dass sich Schülerinnen und Schüler systematisch und mindestens mehrere Monate mit dem Erlernen eines Musikinstruments beschäftigt haben. Bei der Auswertung der Daten für *KoMus* wurde unterschieden zwischen Schülern, die kein Instrument spielen, solchen, die ein bis zwei Jahre ein Instrument gespielt haben und solche, die mehr als zwei Jahre Instrumentalunterricht hatten.

eines 7%-Signifikanzniveaus gesprochen werden; eine DIF-Analyse der Aufgabe ergab keinen signifikanten Unterschied in der Lösungswahrscheinlichkeit zwischen diesen beiden Gruppen. Darüber hinaus haben wir für die Interpretation auch Ergebnisse musikpsychologischer Forschung herangezogen, die ja schon bei der Erstellung des Testkonstrukts eine Rolle spielten. Im vorliegenden Kontext wurde insbesondere die Theorie des Arbeitsgedächtnisses nach Baddeley bedeutsam. Wichtige Begriffe in diesem Kontext sind das Speichermodul ‚phonetische Schleife‘ und so genannte ‚rehearsal-Prozesse‘, was das innerliche Wiederholen dessen bezeichnet, was man sich merken möchte. Betrachtet man die Daten zusammen mit dem theoretischen Hintergrund, so lassen sich mit aller gebotenen Vorsicht folgende Aussagen treffen:

- Die Aufgabe ist für Schülerinnen und Schüler (ohne vorherige intensive Schulung in Gehörbildung) vermutlich nur mit Hilfe der phonetischen Schleife und von rehearsal-Prozessen erfolgreich zu lösen. Rehearsal-Prozesse sind nötig, weil die Kapazität der phonetischen Schleife nicht ausreicht, vier verschiedene Rhythmen abzuspeichern. Das beschreiben die Schüler deutlich im Rahmen der ‚Cognitive Labs‘.
- Diese Strategie wird vermutlich von Schülern mit und ohne Instrumentalunterricht angewendet. Schülern mit intensiver instrumentaler Vorbildung gelingt erwartungsgemäß die Nutzung dieser Strategie etwas erfolgreicher.
- Andere Strategien, die im Rahmen der ‚Cognitive Labs‘ beschrieben wurden, wie beispielsweise das Abzählen von Schlägen, führten nicht zur richtigen Lösung.

Der explorative Einsatz von ‚Cognitive Labs‘ im Kontext des Projekts *KoMus* ermöglichte detaillierten Aufschluss über Lösungsstrategien: Für diese Art von Aufgabe stellte sich heraus, dass mit einer Veränderung der Reihenfolge der Antwortalternativen der Schwierigkeitsgrad der Aufgabe deutlich vermindert oder gesteigert werden könnte. Hier passen die Befunde der ‚Cognitive Labs‘ zu denen, die bei der Analyse der Items erzielt wurden: Die Anordnung der Distraktoren spielt eine wichtige Rolle (s. Abschnitt 4.2.1). Erkenntnisse wurden auch in Bezug auf die getestete Kompetenz gewonnen: Mit Hilfe dieser Aufgabe wird nicht etwa das rationale Erfassen und verbale Kodieren einer rhythmischen Gestalt getestet, sondern die erfolgreiche Nutzung des Arbeitsgedächtnisses. Die hier gewonnenen Ergebnisse passen exakt zur Beschreibung der mittleren Kompetenzausprägung, die bei der Item-Analyse entwickelt wurde, und präzisieren sie in aufschlussreicher Weise. Auch wenn nur eine relativ isolierte Kompetenzfacette näher beleuchtet wer-

den konnte, lässt sich an diesem Beispiel doch schon erkennen, dass eine genauere Kenntnis der Lösungsstrategien nicht nur Erkenntnisse über Kompetenzausprägungen, sondern auch Hinweise auf gezielte Fördermöglichkeiten bereitstellt.

Als Fazit kann festgehalten werden, dass die Methode der ‚Cognitive Labs‘ die Validierung eines Kompetenzmodells, wie sie im Rahmen von *KoMus* angestrebt wird, sinnvoll unterstützen kann, dass sie aber auch darüber hinaus, insbesondere wenn sie vor dem Hintergrund eines qualitativen Paradigmas durchgeführt wird, für musikpädagogische Forschung interessante Perspektiven eröffnet.¹⁷

17 Nur erwähnt werden soll hier die Chance, die der Einsatz von ‚Cognitive Labs‘ in der Ausbildung von Musiklehrern eröffnet: Studierende erfahren im engen Kontakt mit den Schülerinnen und Schülern von deren Lernproblemen und -schwierigkeiten sowie von Strategien, deren Kenntnis für die Gestaltung des eigenen Unterrichts hilfreich sein kann.

Anhang

Tabelle 2

Zusammenfassung der wichtigsten psychometrischen Kennwerte der selektierten Items (Itempool); entnommen aus: Knigge 2010, Kap. 6.1.3

Schwierigkeit (klassisch) = Schwierigkeitsindizes P_i ; Schwierigkeit (Rasch) = Itemparameter des Rasch-Modells; Itemfit (MNSQ) = ‚weighted mean square‘ (Fitmaß des Statistikprogramms ConQuest); MW = Mittelwert; SD = Standardabweichung

- 1 Das EAP/PV-Reliabilitätsmaß (EAP = expected a posteriori, PV = plausible values) ist dem häufig in der Klassischen Testtheorie verwendeten Cronbachs α vergleichbar und führt meist auch zu sehr ähnlichen Resultaten (vgl. Rost, 2004, S. 382).
- 2 In Testheft 6 wurde ein Item trotz einer zu niedrigen Trennschärfe (.19) selektiert. Das Item wies in Bezug auf alle weiteren psychometrischen Kriterien gute Werte auf und war aus inhaltlicher Sicht unverzichtbar.

Testheft Nr.	Itemanzahl insg./selektiert	Schwierigkeit (klassisch)		Schwierigkeit (Rasch)		Itemfit (MNSQ)		Trennschärfe		Reliabilität (EAP/PV) ¹
		Min/Max	MW (SD)	Min/Max	Min/Max	MW (SD)	Min/Max	MW (SD)		
1	20/15	12.71/68.36	39.58 (15.11)	-1.48/2.03	0.88/1.13	1.01 (0.07)	0.32/0.57	0.42 (0.08)	0.693	
2	28/15	5.98/91.45	56.35 (24.60)	-2.43/2.10	0.81/1.15	0.99 (0.11)	0.30/0.65	0.45 (0.12)	0.797	
3	26/15	23.68/94.24	59.36 (25.61)	-2.42/2.25	0.88/1.10	1.00 (0.06)	0.28/0.57	0.41 (0.10)	0.720	
4	32/23	14.88/94.69	62.55 (30.40)	-2.20/3.18	0.83/1.11	1.02 (0.05)	0.26/0.52	0.37 (0.07)	0.738	
5	41/29	15.77/90.04	57.61 (22.23)	-2.34/2.57	0.88/1.15	1.00 (0.07)	0.25/0.56	0.39 (0.09)	0.822	
6	31/24	6.67/87.50	36.19 (24.26)	-2.46/2.54	0.89/1.08	1.00 (0.05)	0.19/0.61 ²	0.36 (0.10)	0.719	
7	27/15	5.23/66.56	29.66 (19.54)	-1.67/2.42	0.94/1.11	1.01 (0.05)	0.28/0.53	0.39 (0.09)	0.574	
8	34/24	6.20/90.08	33.39 (22.71)	-3.55/2.49	0.81/1.19	1.00 (0.09)	0.26/0.55	0.42 (0.08)	0.823	
9	35/19	7.97/88.05	34.78 (23.00)	-2.83/1.48	0.86/1.07	1.00 (0.05)	0.27/0.53	0.37 (0.09)	0.632	
Gesamt	275/179									

Literatur

- Benner, D. (2007): Unterricht - Wissen - Kompetenz. Zur Differenz zwischen didaktischen Aufgaben und Testaufgaben. In: Benner, D. (Hg.): Bildungsstandards. Instrumente zur Qualitätssicherung im Bildungswesen. Kontroversen - Beispiele - Perspektiven. Paderborn: Schöningh, S. 124–138.
- Bruhn, H. (2005): Wissen und Gedächtnis. In: Oerter, R.; Stoffer, T. H. (Hg.): Allgemeine Musikpsychologie. Göttingen: Hogrefe (Enzyklopädie der Psychologie, Serie VII, Bd. 1), S. 537–590.
- Büchter, A.; Leuders, T. (2005): Mathematikaufgaben selbst entwickeln. Lernen fördern - Leistung überprüfen. Berlin: Cornelsen Scriptor.
- Bühner, M. (2006): Einführung in die Test- und Fragebogenkonstruktion. 2., aktualisierte und erw. Auflage. München: Pearson Studium.
- Caspari, D.; Grotjahn, R.; Kleppin, K. (2008): Kompetenzorientierung und Aufgaben. Zur Unterscheidung zwischen Lern- und Testaufgaben. In: Tesch, B.; Leupold, E.; Köller, O. (Hg.): Bildungsstandards Französisch: konkret. Sekundarstufe I: Grundlagen, Aufgabenbeispiele und Unterrichtsanregungen. Berlin: Cornelsen Scriptor, S. 85–87.
- Criblez, L.; Oelkers, J.; Reusser, K.; Berner, E.; Halbheer, U.; Huber, C. (2009): Bildungsstandards. Zug: Klett und Balmer (Lehren lernen - Basiswissen für die Lehrerinnen- und Lehrerbildung).
- Cvetko, A.; Meyer, D. (2009): Problemlösen im Musikunterricht – Interdisziplinarität als Ausgangspunkt für eine kompetenzorientierte Perspektive. In: Schläbitz, N. (Hg.): Interdisziplinarität als Herausforderung musikpädagogischer Forschung. Essen: Die Blaue Eule (Musikpädagogische Forschung, Bd. 30).
- Ericsson, K. A.; Simon, H. A. (1999): Protocol analysis. Verbal reports as data. 3. überarbeitete Auflage. Cambridge: Mit Press.
- Gembris, H. (2005): Die Entwicklung musikalischer Fähigkeiten. In: La Motte-Haber, H. de; Rötter, G. (Hg.): Musikpsychologie. Laaber: Laaber (Handbuch der Systematischen Musikwissenschaft, 3), S. 394–456.
- Hartig, J. (2007): Skalierung und Definition von Kompetenzniveaus. In: Beck, B.; Klieme, E. (Hg.): Sprachliche Kompetenzen - Konzepte und Messung. DESI-Studie (Deutsch Englisch Schülerleistungen International). Weinheim: Beltz (DESI Ergebnisse, 1), S. 83–99.

- Hartig, J.; Jude, N. (2007): Empirische Erfassung von Kompetenzen und psychometrische Kompetenzmodelle. In: Hartig, J.; Klieme, E. (Hg.): Möglichkeiten und Voraussetzungen technologiebasierter Kompetenzdiagnostik. Eine Expertise im Auftrag des Bundesministeriums für Bildung und Forschung. Berlin: BMBF (Bildungsforschung, 20), S. 17–36.
- Jank, W. (Hg.) (2007): Musik-Didaktik. Praxishandbuch für die Sekundarstufe I und II. 2. Auflage. Berlin: Cornelsen Scriptor.
- Jordan, A.-K., Knigge, J., Lehmann-Wermser, A., Lehmann, A. C. & Niessen, A. (i. Vorb.). Entwicklung und Validierung eines Kompetenzmodells im Fach Musik – Wahrnehmen und Kontextualisieren von Musik.
- Jordan, A.-K.; Lehmann, A. C.; Knigge, J. (2010): IRT-basierte Kompetenzmodellierung – Erste Ergebnisse der Validierung eines Kompetenzmodells für den Bereich „Musik wahrnehmen und kontextualisieren“. In: Knolle, N. (Hg.): Evaluationsforschung in der Musikpädagogik. Essen: Die Blaue Eule (Musikpädagogische Forschung, 31).
- Klieme, E.; Avenarius, H.; Blum, W., et al. (Hg.) (2003): Zur Entwicklung nationaler Bildungsstandards. Eine Expertise. Berlin: BMBF (Bildungsforschung, 1).
- Klieme, E.; Hartig, J. (2007): Kompetenzkonzepte in den Sozialwissenschaften und im erziehungswissenschaftlichen Diskurs. In: Prenzel, M.; Gogolin, I.; Krüger, H. (Hg.): Kompetenzdiagnostik (Zeitschrift für Erziehungswissenschaft Sonderheft, 8). Wiesbaden: VS Verlag für Sozialwissenschaften, S. 11–29.
- Knigge, J. (2008): Hinweise zur Erstellung von Testaufgaben für das KoMus-Projekt. (unveröffentlichtes Papier). Universität Bremen, Institut für Musikwissenschaft und Musikpädagogik – Projekt KoMus.
- Knigge, J. (2010): Modellbasierte Entwicklung und Analyse von Testaufgaben zur Erfassung der Kompetenz „Musik wahrnehmen und kontextualisieren“. Dissertation. Universität Bremen.
- Knigge, J.; Lehmann-Wermser, A. (2008): Bildungsstandards für das Fach Musik - Eine Zwischenbilanz. In: Zeitschrift für Kritische Musikpädagogik, Sonderedition: Bildungsstandards und Kompetenzmodelle für das Fach Musik? S. 60–98. Online verfügbar unter <http://www.zfkm.org/sonder08-knigge-lehmannwermser.pdf>, zuletzt geprüft am 27.08.2009.

- Köller, O. (2008): Bildungsstandards – Verfahren und Kriterien bei der Entwicklung von Messinstrumenten. In: Zeitschrift für Pädagogik, Jg. 54, H. 2, S. 163–173.
- Köller, O.; Böhme, K.; Winkelmann, H.; Bremerich-Vos, A.; Granzer, D.; Vock, M.; Pöhlmann, C.; Robitzsch, A.; Würfel, K. (2005): Hinweise zur Erstellung von Testaufgaben für das Projekt "Evaluation der Standards Deutsch in der Grundschule" ESDeG (Primarbereich, Jahrgang 4). (unveröffentlichtes Papier). IQB. Berlin.
- Kreutz, G. (2005): Melodiewahrnehmung: Funktionen von Arbeitsgedächtnis und Aufmerksamkeit. In: La Motte-Haber, H. de; Rötter, G. (Hg.): Musikpsychologie. Laaber: Laaber (Handbuch der Systematischen Musikwissenschaft, 3), S. 185–207.
- La Motte-Haber, H. de (2005): Modelle der musikalischen Wahrnehmung. Psychophysik - Gestalt - Invarianten - Mustererkennen - Neuronale Netze - Sprachmetapher. In: La Motte-Haber, H. de; Rötter, G. (Hg.): Musikpsychologie. Laaber: Laaber (Handbuch der Systematischen Musikwissenschaft, 3), S. 55–73.
- Lange, E. B. (2005): Musikpsychologische Forschung im Kontext allgemeinspsychologischer Gedächtnismodelle. In: La Motte-Haber, H. de; Rötter, G. (Hg.): Musikpsychologie. Laaber: Laaber (Handbuch der Systematischen Musikwissenschaft, 3), S. 74–100.
- Mislevy, R. J.; Steinberg, L. S.; Almond, R. G. (2002): On the Roles of Task Model Variables in Assessment Design. In: Irvine, S. H. (Hg.): Item generation for test development. Mahwah: Lawrence Erlbaum, S. 97–128.
- Moosbrugger, H.; Kelava, A. (2007): Qualitätsanforderungen an einen psychologischen Test (Testgütekriterien). In: Moosbrugger, H.; Kelava, A. (Hg.): Testtheorie und Fragebogenkonstruktion. Heidelberg: Springer, S. 7–26.
- Nauck-Börner, C. (1987): Wahrnehmung und Gedächtnis. In: La Motte-Haber, H. de (Hg.): Psychologische Grundlagen des Musiklernens. Kassel: Bärenreiter (Handbuch der Musikpädagogik, 4), S. 13–115.
- Niessen, A. (2008): Leistungsmessung oder individuelle Förderung? Zur Funktion und Gestaltung von Aufgaben im Unterricht. In: Schäfer-Lembeck, H.-U. (Hg.): Leistung im Musikunterricht. Beiträge der Münchner Tagung 2008. München: Allitera (Musikpädagogische Schriften der Hochschule für Musik und Theater München, 2), S. 134–152.

- Niessen, A. (2009): Kompetenzlernen und verstehende Auseinandersetzung mit Musik. Eine Antwort auf Christoph Richter. In: Diskussion Musikpädagogik, Nr. 44, S. 58–60.
- Niessen, A.; Lehmann-Wermser, A.; Knigge, J.; Lehmann, A. C. (2008): Entwurf eines Kompetenzmodells 'Musik wahrnehmen und kontextualisieren'. In: Zeitschrift für Kritische Musikpädagogik, Sonderedition: Bildungsstandards und Kompetenzmodelle für das Fach Musik?, S. 3–33. Online verfügbar unter <http://www.zfkm.org/sonder08-niessenetal.pdf>, zuletzt geprüft am 27.08.2009.
- Nold, G.; Rossa, H. (2007): Hörverstehen. In: Beck, B.; Klieme, E. (Hg.): Sprachliche Kompetenzen - Konzepte und Messung. DESI-Studie (Deutsch Englisch Schülerleistungen International). Weinheim: Beltz (DESI Ergebnisse, 1), S. 178–196.
- Prenzel, M.; Häußler, P.; Rost, J.; Senkbeil, M. (2002): Der PISA-Naturwissenschaftstest: Lassen sich die Aufgabenschwierigkeiten vorher-sagen? In: Unterrichtswissenschaft, Jg. 30, H. 2, S. 120–135.
- Richter, C. (2008): Musikunterricht "von unten". Curriculare Arbeit und aufbauender Unterricht von den Schülern aus. In: Diskussion Musikpädagogik, Nr. 37, S. 11–20.
- Richter, C. (2009): Musikunterricht am Scheidewege. Kompetenzlernen oder verstehende Auseinandersetzung mit Musik. In: Diskussion Musikpädagogik, Nr. 42, S. 7–8.
- Rost, J. (2004): Lehrbuch Testtheorie - Testkonstruktion. 2., vollst. überarb. und erw. Aufl. Bern: Huber.
- Runfola, M.; Swanwick, K. (2002): Developmental characteristics of music learners. In: Colwell, R.; Richardson, C. P. (Hg.): The new handbook of research on music teaching and learning. A project of the Music Educators National Conference. New York: Oxford Univ. Press, S. 373–397.
- Stoffer, T. H. (2005): Aufmerksamkeitsprozesse beim Musikhören: Wissensunabhängige und wissensabhängige Selektionsprozesse. In: Oerter, R.; Stoffer, T. H. (Hg.): Allgemeine Musikpsychologie. Göttingen: Hogrefe (Enzyklopädie der Psychologie, Serie VII, Bd. 1), S. 591–656.
- Strauss, A. L. (1994): Grundlagen qualitativer Sozialforschung. Datenanalyse und Theoriebildung in der empirischen soziologischen Forschung. München: Fink.

Kompetenzmodellierung mit Methoden der Item-Response-Theorie (IRT)

Erste Ergebnisse der Validierung eines Modells für den Bereich „Musik wahrnehmen und kontextualisieren“

1 Einleitung

Seit etwa 10 Jahren gibt es einen Trend in der empirischen Bildungsforschung, der maßgeblich durch die Diskussionen um die Ergebnisse der PISA-Studien angestoßen wurde. Seither haben Bildungsforscher das Konzept der Kompetenzen zu einem zentralen Thema gemacht und damit eine Abkehr in der Betrachtung vollzogen von einem traditionell am Input orientierten Verständnis von Bildungsvorgängen hin zu einer am Output orientierten Sichtweise. Dieser Paradigmenwechsel hat in vielen Didaktikfächern tief greifende und sichtbare Veränderungen nach sich gezogen. Auch in der Musikpädagogik sind in den letzten Jahren zum Teil kontroverse Diskussionen geführt worden (Bähr 2004, Richter 2007). Diese Auseinandersetzung begreift sich als eher konzeptionell, und es geht ihr weniger um den Anschluss an Bemühungen der empirischen Erforschung, wie sie in den Erziehungswissenschaften derzeit üblich sind.

Im Rahmen der Diskussion um Bildungsstandards und Kompetenzorientierung kommt der Modellierung von Kompetenzen eine besondere Rolle zu. In dem musikpädagogischen Projekt KoMus, von dem hier die Rede sein soll, ging es darum, zunächst einmal auf der Grundlage von Überlegungen aus der Fachdidaktik heraus ein vorläufiges theoretisches Modell zu entwickeln, wie weiter unten näher beschrieben wird. Dabei fanden neben der Sichtung curricularen Materials (Knigge, Lehmann-Wermser 2008) auch Forschungsergebnisse der Musikpädagogik und Musikpsychologie Berücksichtigung (z. B. Bruhn 2005, La Motte-Haber 2005, Lange 2005, Serafine 1988, Stoffer 2005). Das resultierende vorläufige Modell wurde dann als Konstrukt zum Zweck der späteren empirischen Überprüfung mit Hilfe von Testaufgaben operationalisiert. Wie diese Operationalisierung im Projekt KoMus erfolgte, wird an ande-

rer Stelle beschrieben (Knigge, Niessen & Jordan 2010). Nach der aufwändigen Entwicklung von Aufgaben für eine bestimmte Zielgruppe (Schüler der Klassenstufe sechs) konnten diese als Test zusammengefasst einer größeren Anzahl von Schülern vorgelegt werden. Die dabei anfallenden Daten waren die Grundlage für eine Modellierung des Kompetenzmodells, um die es in diesem Beitrag gehen wird.

Während es in den schulischen ‚Kernfächern‘ Mathematik, Deutsch und Englisch bereits groß angelegte Studien (‚large-scale assessments‘) zu den von Schülerinnen und Schülern¹ erworbenen Kompetenzen gibt (z. B. TIMSS, DESI, IGLU), liegen für die deutsche Musikpädagogik erst seit etwa drei Jahren Überlegungen zur empirischen Erforschung von Kompetenzen im Fach Musik vor (Niessen et al. 2008). Hierbei wurde mit Methoden der Item-Response-Theorie gearbeitet (IRT, s. u. für weitere Details), während in der Vergangenheit Leistungstests auf der Basis der klassischen Testtheorie (KTT) ausgewertet wurden (z. B. Lohmann 1982, Bähr 2001). Im Vordergrund des vorliegenden Beitrages steht die empirische Validierung des theoretischen Kompetenzmodells durch Überprüfung seiner dimensionalen Struktur. Einleitend soll kurz dargestellt werden, was allgemein unter Kompetenzmodellen verstanden wird und welche spezifischen Formen von Kompetenzmodellen differenziert werden können.

Im Zentrum unserer Ausführungen stehen die empirische Validierung des theoretischen Kompetenzmodells und damit die Überprüfung seiner dimensional Struktur. Hierbei wird hinterfragt, ob die von uns im Test angelegten inhaltlichen Dimensionen auch empirisch nachweisbar sind. Bevor wir einige Ergebnisse aus dem KoMus-Projekt berichten, wollen wir jedoch zum besseren Verständnis beschreiben, was allgemein unter Kompetenzmodellen verstanden wird und welche spezifischen Formen von Kompetenzmodellen es gibt.

Kompetenzmodelle

Auch wenn eine ausführliche Diskussion des Kompetenzbegriffs hier nicht erfolgen kann, sei darauf verwiesen, dass dem vorliegenden Projekt der Kompetenzbegriff nach Weinert (2001) zugrunde liegt, wie er auch im aktuellen erziehungswissenschaftlichen und pädagogisch-psychologischen Diskurs Ver-

1 Aus Gründen der leichteren Lesbarkeit verzichten wir im Folgenden auf die Nennung beider Geschlechter.

wendung findet (Klieme et al. 2003, vgl. auch Knigge & Lehmann-Wermser 2009). Kompetenzen sind danach „die bei Individuen verfügbaren oder durch sie erlernbaren kognitiven Fähigkeiten und Fertigkeiten, um bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, um die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können“ (Weinert 2001, S. 27). Ein Kompetenzmodell unterscheidet „Teildimensionen innerhalb einer Domäne [...] und es beschreibt jeweils unterschiedliche Niveaustufen auf solchen Dimensionen. Jede Kompetenzstufe ist durch kognitive Prozesse und Handlungen von bestimmter Qualität spezifiziert, die Schüler auf dieser Stufe bewältigen können, nicht aber Schüler auf niedrigeren Stufen“ (Klieme et al. 2003, S. 24). Jede dieser Teildimensionen, z. B. die Fähigkeit des Heraushörens eines Instruments, kann durch verschiedene Niveaustufen zunehmenden Anspruchs beschrieben werden. Im Hinblick auf das Heraushören eines Instruments hieße das etwa, ein Solo-Instrument an seinem Klang zu erkennen bzw. einzelne Instrumente aus einem Orchesterklang herauszuhören. Um eine solche Beschreibung vornehmen zu können, muss das zunächst theoretische Modell aber operationalisiert werden, d. h., es müssen Aufgaben entwickelt werden, die die Aspekte des Modells konkretisieren (vgl. Knigge et al. 2010).

Ziel eines Kompetenzmodells ist erstens die Beschreibung von Anforderungen, „deren Bewältigung von Schülerinnen und Schülern erwartet wird“ und zweitens die Entwicklung begründeter Vorstellungen darüber, „welche Abstufungen eine Kompetenz annehmen kann bzw. welche Grade oder Niveaustufen sich bei den einzelnen Schülerinnen und Schülern feststellen lassen“ (Klieme et al. 2003, S. 74). „Das Erreichen einer Kompetenzstufe sagt etwas darüber aus, welche Handlungen und mentalen Operationen mit hoher Wahrscheinlichkeit korrekt ausgeführt werden können“ (Klieme et al. 2003, S. 23). Kompetenzmodelle sind also nicht gleichzusetzen mit Leistungstests im traditionellen Sinn; vielmehr sind sie wissenschaftliche Konstrukte.

Zusammenfassend stellen Kompetenzmodelle eine wichtige vermittelnde Schnittstelle zwischen verallgemeinerten Zielen von Unterricht und Testaufgaben dar. Sie geben Hinweise auf die Aufgabenkonstruktion, „indem sie kognitive Leistungen mit unterschiedlichem Schwierigkeitsniveau spezifizieren. Umgekehrt ermöglicht erst die Einordnung in ein Kompetenzmodell, zu verstehen, was das Lösen oder Nichtlösen einer Aufgabe bedeutet [...]. Modellvorstellungen, die den Aufbau von Kompetenzen über mehrere Stufen hinweg charakterisieren, sind demnach wichtige Orientierungen für die Unter-

richtspraxis und die Bewertung von Lernergebnissen“ (Klieme et al. 2003, S. 24).

Formen von Kompetenzmodellen

In der empirischen Bildungsforschung unterscheidet man zwei Formen von Kompetenzmodellen: *Kompetenzstruktur-* und *Kompetenzniveaumodelle* (Hartig, Klieme 2006; vgl. auch Jordan, Knigge i. Druck). Erstere beschreiben die Grundstruktur bzw. den Aufbau einer Kompetenz ausgehend von den statistischen Zusammenhängen zwischen den auf ein Modell bezogenen Skalen und Tests (Hartig, Klieme 2006; s. auch Abb. 5). Mithilfe faktorenanalytischer Verfahren werden Aufgaben, die hohe Zusammenhänge untereinander zeigen, zu Dimensionen zusammengefasst. Man geht somit davon aus, dass die Aufgaben dasselbe Merkmal erfassen (Hartig, Klieme 2006). Bei der Entscheidung für ein bestimmtes Strukturmodell (Beispiel 3-dimensionales Modell s. linker Kasten Abb. 1), also der Frage, wie viele Dimensionen man differenzieren möchte und wie viele sinnvoll sind, muss jeweils eine Abwägung pragmatischer und theoretischer Aspekte vorgenommen werden. Kompetenzstrukturmodelle können sich auch, statt mit der Kompetenz als solcher, mit der Binnenstruktur einzelner Kompetenzbereiche befassen, d. h. mit den zugrunde liegenden Teilkompetenzen und ihren Zusammenhängen (Hartig, Klieme 2007; s. auch rechter Kasten in Abb. 1).

Dieser Betrachtung liegt die Einsicht zugrunde, dass Kompetenzstrukturmodelle eine ein- oder mehrdimensionale Struktur aufweisen können. Ein eindimensionales Modell beschreibt die Unterschiede in der zu erfassenden Kompetenz auf einem einzelnen Kontinuum. Um differenziertere Aussagen über Kompetenzen und Teilkompetenzen der untersuchten Schüler machen zu können und um komplexere Kompetenzstrukturen abzubilden, können auch mehrdimensionale Item-Response-Modelle verwendet werden. Nach der Identifizierung der Kompetenzstruktur kann mithilfe von *Kompetenzniveaumodellen* erfasst werden, welche fachbezogenen Leistungsanforderungen bzw. welche spezifischen Kompetenzen Schüler bewältigen können und welche nicht.

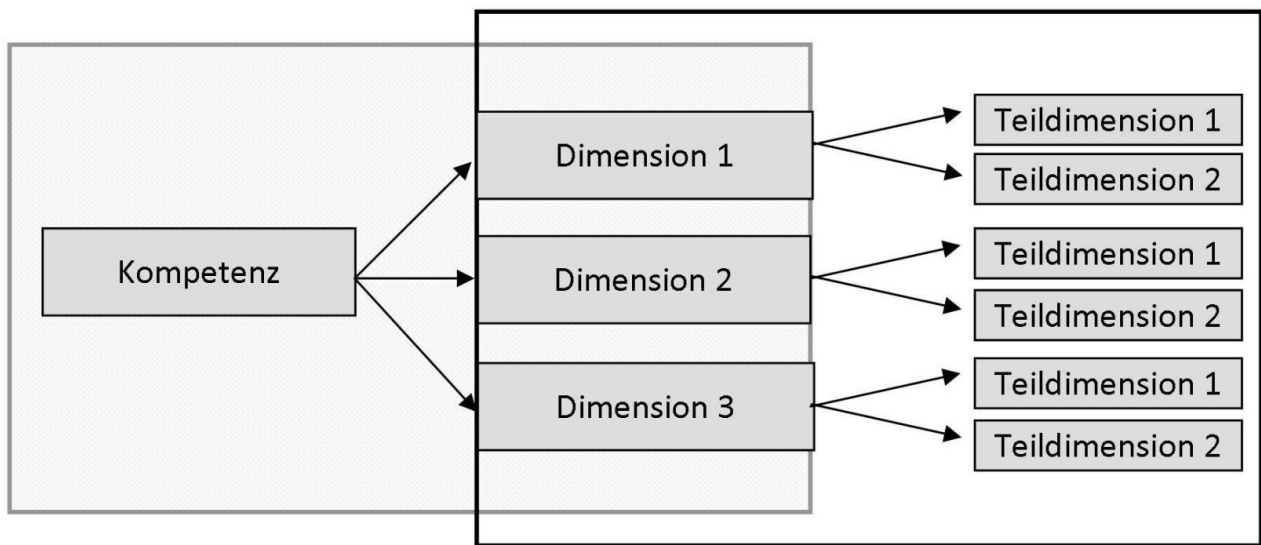


Abb. 1: Schematische Darstellung eines Kompetenzstrukturmodells (Erläuterung s. Text)

Betrachten wir zunächst ein eindimensionales Kompetenzkonstrukt, wobei alle Leistungen auf einem Kontinuum abgebildet werden können, z. B. verschiedene Leistungsniveaus der Wahrnehmungsfähigkeit von Schülern. Dieses Kontinuum wird in Abschnitte unterteilt, welche als *Kompetenzniveaus* oder auch *Kompetenzstufen* bezeichnet werden. In Anlehnung an Hartig und Klie-me (2006) soll im Folgenden die Bezeichnung Kompetenzniveaus verwendet werden. Für diese Kompetenzniveaus wird dann eine kriteriumsorientierte Beschreibung der erfassten Kompetenzen durchgeführt, d. h. es wird ein vorab definiertes *Kriterium* der (quantifizierten) Leistungswerte herangezogen. Beispielsweise wäre das einzelne Erkennen aus einem Ensembleklang Niveau II und das gleichzeitige Verfolgen mehrerer Stimmen in einem Musikstück Niveau III (vgl. Abb. 2).

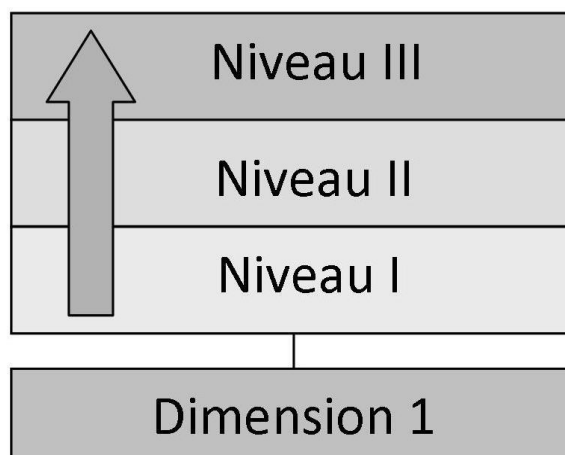


Abb. 2: Schematische Darstellung eines Kompetenzniveaumodells mit einer Dimension

Methoden der Item-Response-Theorie (IRT) (z. B. Rost 2004; Hambleton, Swaminathan 2000) ermöglichen das Bilden einer gemeinsamen Skala, auf der sowohl die Fähigkeiten der Schüler als auch die Schwierigkeiten der Aufgaben dargestellt werden können. Somit können Aussagen über Schüler mit unterschiedlich ausgeprägten Kompetenzen gemacht werden. Gleichzeitig kann festgestellt werden, welche Aufgaben diese Schüler mit einer bestimmten Wahrscheinlichkeit lösen können und welche nicht. Im Gegensatz zur klassischen Testtheorie besteht bei der Item-Response-Theorie die Möglichkeit, Personenfähigkeit und Itemschwierigkeit auf einer gemeinsamen Skala darzustellen. Es kann gezeigt werden, wie sich die Schwierigkeit einzelner Items über das gesamte Modell verteilt. Man kann daher z. B. sagen: „Der Schüler hat eine Fähigkeit, die es ihm erlaubt, die Aufgaben einer ganz bestimmten Schwierigkeitsstufe mit hoher Wahrscheinlichkeit zu lösen, z. B. solche, bei denen sowohl Violin- als auch Bassschlüssel verwendet wird.“ Interessant wird es, wenn man dann auch unterschiedliche Kompetenzbereiche (z. B. Notenlesen, Melodien merken) auf der gleichen Metrik abbilden kann. Kompetenzniveau- und Kompetenzstrukturmodelle schließen sich keineswegs aus, sondern ergänzen sich idealerweise.

Theoretisch entwickeltes Kompetenzmodell für den Bereich „Musik wahrnehmen und kontextualisieren“

Bevor eine Modellierung der Kompetenzen vorgenommen werden kann, müssen auf Basis fachdidaktischen Wissens sowohl ein theoretisches Kompetenzmodell als auch damit verbundene Testaufgaben erstellt werden. Das im Rahmen von KoMus entwickelte Modell bzw. die darauf basierenden Testaufgaben beziehen sich auf vier Dimensionen: reine Wahrnehmung (1), Wahrnehmung in Verbindung mit Fachterminologie (2), in Verbindung mit Notation (3) und in Verbindung mit der Anwendung von Kontextwissen (4) (s. Abb. 3). Eine ausführliche konzeptionelle und inhaltliche Erläuterung der einzelnen Dimensionen ist bei Niessen et al. 2008 nachzulesen. Nach der Darlegung der theoretischen Grundlagen von Kompetenzmodellierung wenden wir uns dem von uns erhobenen Datensatz zu.

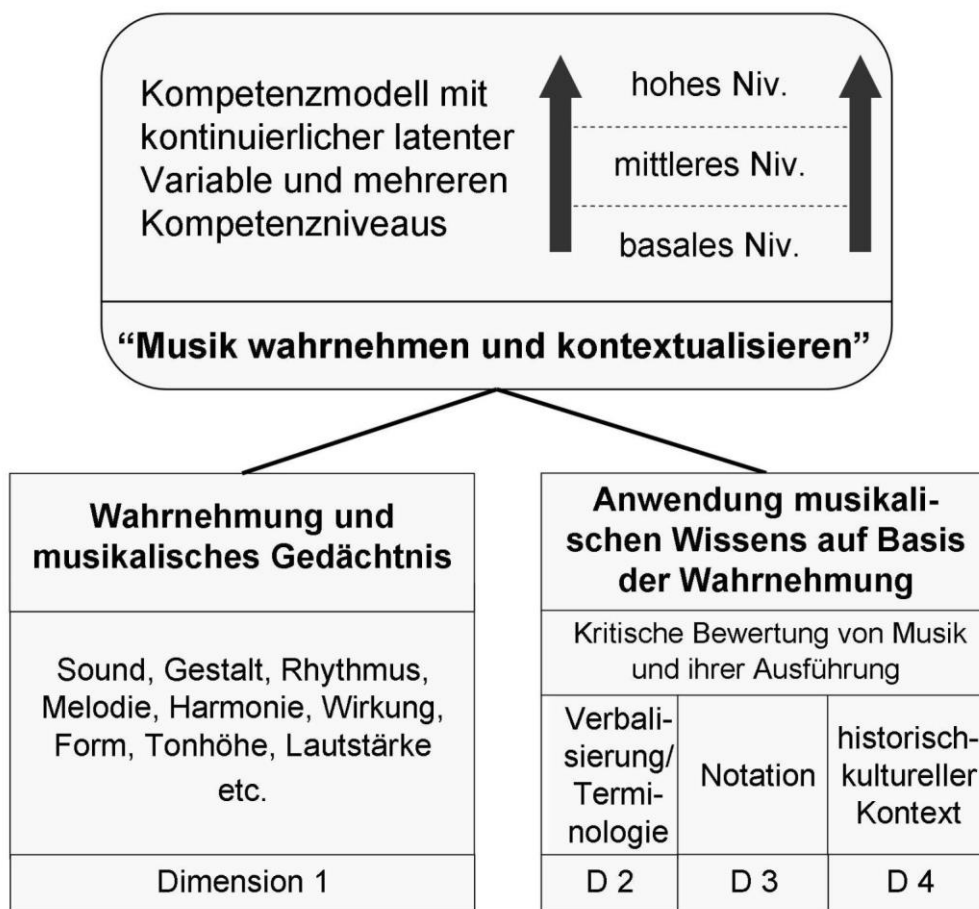


Abb. 3: Kompetenzmodell von Niessen et al. (2008) am Ende der Aufgabenentwicklung².

2 Forschungsdesign und Datengrundlage

Um den Forschungsprozess vom theoretisch entwickelten Kompetenzmodell bis zur empirischen Überprüfung nachvollziehbar machen zu können, werden wir zunächst das Forschungsdesign der Pilotierungsstudie vorstellen. Der Prozess der Aufgabenentwicklung und die damit verbundenen statistischen Analysen werden an anderer Stelle thematisiert (Knigge, Niessen & Jordan 2010).

Da im Fach Musik bisher noch keine Kompetenzmodelle entwickelt wurden, ist es in einem ersten Schritt notwendig, die inhaltlich angelegten und

² Abbildung 3 zeigt eine überarbeitete Fassung des bei Niessen et al. 2008 vorgestellten Modells. In vorliegender Form diente das Modell als Grundlage für die Operationalisierung in Form von Testaufgaben. Die endgültige Formulierung des Modells ist erst nach Abschluss der Analysen der Pilotierungsstudie vorgesehen, sodass das abgebildete Modell weiterhin einen vorläufigen Status besitzt.

vermuteten Dimensionen des Modells empirisch zu validieren. Erst im Anschluss daran kann eine Gesamtmodellierung aller Dimensionen vorgenommen werden. Modellierung heißt in diesem Zusammenhang die Berechnung und Beurteilung der Modellstruktur anhand statistischer Kennwerte.

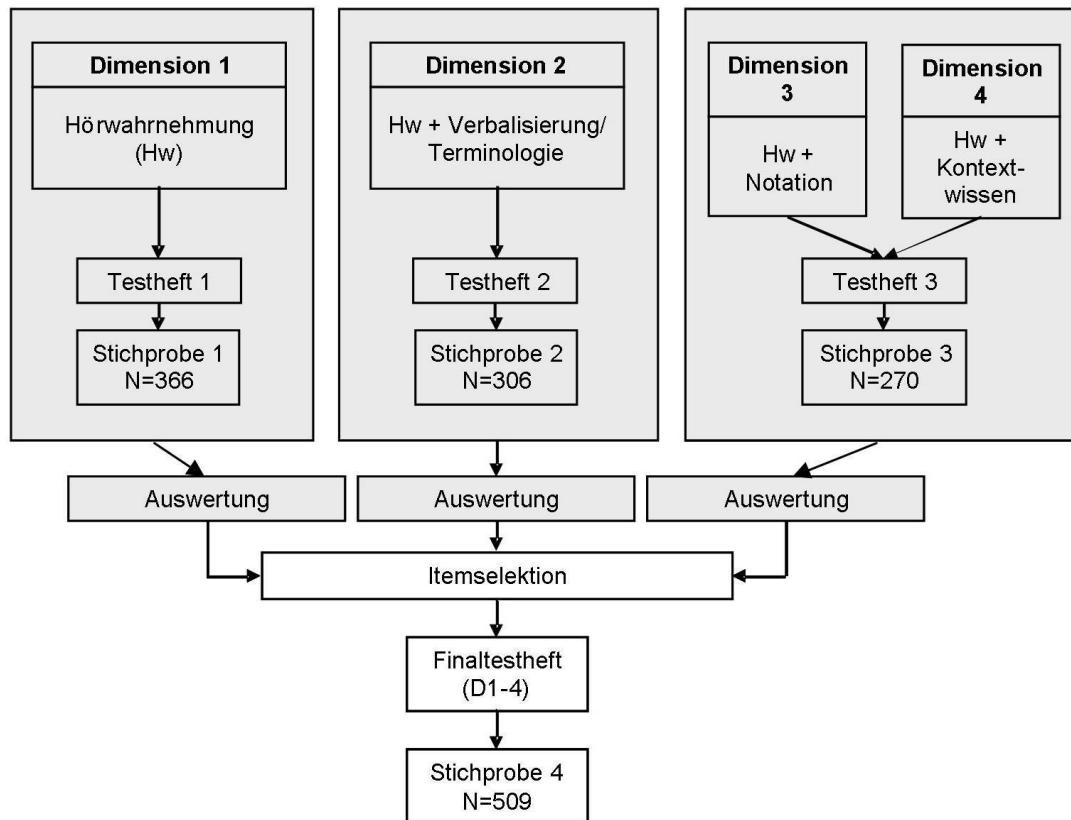


Abb. 4: Testdesign der Pilotierungsstudie

Wie bereits erläutert wurde ein zweistufiges Pilotierungsdesign gewählt. Zunächst wurde pro Dimension ein Testheft mit Aufgaben zusammengestellt und mit einer Schülerstichprobe bestehend aus Schülern von sechsten Klassen in Bremen und Niedersachsen ($N_{(\text{Dim } 1)}=366$, $N_{(\text{Dim } 2)}=306$, $N_{(\text{Dim } 3-4)}^3=270$) getestet. Innerhalb des Klassenverbandes bearbeitete jeder Schüler die Testaufgaben an einem eigenen Laptop und konnte die Hörbeispiele individuell mittels Kopfhörer abspielen. Dabei wurde ein so genanntes Ankeritemdesign verwendet. Eine bestimmte Anzahl von gemeinsamen Items, sog. Ankeritems, wurde von allen Schülern bearbeitet. Diese gemeinsamen Items verbinden („verankern“) die Testhefte. Zusätzlich bearbeiteten die Schüler jeweils weite-

3 Aufgrund der geringer Itemanzahl in Dimension 3 und 4 wurden die Aufgaben dieser beiden Dimensionen zusammen in einem Testheft bearbeitet.

re Items. Der Vorteil ist hierbei, dass so eine größere Anzahl von Items einbezogen werden kann, als dies möglich wäre, wenn jede Testperson die gleichen Testitems vorgelegt bekäme. Über die Ankeritems kann später mithilfe der Item-Response-Theorie eine Skalierung für alle Aufgaben vorgenommen werden, auch wenn diese nicht von allen Schülern bearbeitet wurden.

Nach der Auswertung der Testhefte 1 bis 3 wurden Aufgaben für das Finaltestheft ausgewählt. Bei der Auswahl wurde neben der Berücksichtigung statistischer Kriterien (Trennschärfe, Itemschwierigkeit, Modellfit, etc.) auch darauf geachtet, dass die Aufgaben hinsichtlich ihrer Schwierigkeit gut über das gesamte Kontinuum verteilt und alle inhaltlichen Facetten des Modells vertreten waren. Im Finaltestheft wurden also qualitativ hochwertige Items aus allen Dimensionen einbezogen. Da aufgrund mangelnder Testzeit nicht alle Aufgaben der Dimensionen 1 bis 4 in den Finaltest übernommen werden konnten, mussten für jede Dimension je neun Ankeritems für das Finaltestheft ausgewählt werden. Neben diesen 36 Items wurden 71 weitere Items ausgewählt, die aber nicht von allen Schülern bearbeitet wurden (Erläuterung s. o.). Das Finaltestheft wurde mit 508 Schülern verschiedener Schulformen aus Bremen und Niedersachsen getestet und ermöglicht nun die Überprüfung von Zusammenhängen zwischen den Dimensionen. Insgesamt wurden somit 1451 Schüler an 27 Schulen verschiedener Schulformen⁴ in die Pilotierung einbezogen.

3 Forschungsleitende Fragestellungen

Mit Hilfe der Daten der Pilotierungsstudie sollen zwei Fragestellungen bearbeitet werden: In einem ersten Schritt wird überprüft, ob das Raschmodell (das einfachste Testmodell der Item-Response-Theorie) gültig ist und auf die Daten angewendet werden kann. Anders ausgedrückt wird untersucht, ob sich die musikalische Kompetenz als latente Variable abbilden lässt. Nach verschiedenen standardmäßigen Modellgültigkeitsüberprüfungen mit Hilfe der Software *ConQuest* kann die Struktur der Daten genauer untersucht werden. Unter der Voraussetzung, dass das Raschmodell gilt, kann mit so genannten Dimensionsanalysen im zweiten Schritt geklärt werden, ob eine eindimensionale oder eine mehrdimensionale Struktur vorliegt. Dabei können verschiedene Modellstrukturen (z. B. zwei- vs. vierdimensionales Modell) einander gegenüberge-

4 12 Gymnasien, 11 Haupt- und Realschulen, 2 Gesamtschulen, 2 Sekundarschulen.

stellt werden und somit das Modell identifiziert werden, das die Daten am besten abbildet.

Nach Aufstellung der Modellstruktur beschäftigen sich die Analysen mit den Zusammenhängen zwischen den gefundenen Dimensionen. Es wird dabei überprüft, ob die Teildimensionen tatsächlich eigenständige Dimensionen bilden oder ob die Korrelationen zwischen den Teildimensionen so hoch sind, dass sie nur in der Zusammenfassung zu einer Dimension sinnvoll interpretierbar sind. Weiterhin wird die interne Konsistenz (Reliabilität) jeder einzelnen Dimension berechnet, die einen Kennwert für ihre Stabilität darstellt.

4 Methoden

Modellierung der Kompetenz „Musik wahrnehmen und kontextualisieren“ als latente Variable - Vorgehensweise

Um die Gültigkeit des Raschmodells zu überprüfen, gibt es eine Reihe von Modellgültigkeitstests, die hier nicht im Einzelnen aufgeführt werden können (s. dazu Rost 2004, Bühner 2008, Moosbrugger et al. 2007). Eine zentrale Annahme des Raschmodells bildet die Annahme der Personenhomogenität, die besagt, dass „alle getesteten Personen den Test aufgrund derselben Eigenschaft oder Fähigkeit bearbeiten“ (Rost 2004, S. 347). Diese Annahme kann mit dem Likelihood-Quotienten-Test (auch Andersen-Test) überprüft werden (Andersen 1973, Bühner 2008). Dabei werden die Itemparameter für verschiedene Untergruppen geschätzt und auf Unterschiede geprüft. Die Untergruppen können nach verschiedenen Teilungskriterien, z. B. Testscore, Geschlecht oder Schulform, gebildet werden. Im Hinblick auf die musikalische Kompetenz ist es darüber hinaus naheliegend, die Stichprobe nach dem Kriterium der vorhandenen bzw. nicht vorhandenen Instrumentalerfahrung zu teilen. Zur besseren Anschaulichkeit können die Itemparameter der verschiedenen Untergruppen in einem Streudiagramm einander gegenübergestellt werden (s. Abb. 6). Bei Gültigkeit des Raschmodells müssten die Itemparameter eine Gerade mit Steigung 1 bilden, die durch den Nullpunkt verläuft, da es keine großen Abweichungen zwischen den Itemparametern der verschiedenen Gruppen (z. B. Schüler mit/ohne Instrumentalerfahrung) geben dürfte.

Unter der Voraussetzung, dass das Raschmodell gilt, schließen sich weitere Analysen an. Es gilt zu hinterfragen, ob die Daten eine ein- oder mehrdimensionale Struktur aufweisen. Zur Untersuchung des Kompetenzstrukturmodells

(s. Abschnitt 1) kann ein exploratives ebenso wie ein konfirmatorisches Vorgehen genutzt werden, vergleichbar mit der Unterscheidung bei der Faktorenanalyse aus der Klassischen Testtheorie (Bühner 2008). Man nimmt an, dass die Interkorrelationen zwischen den Einzelitems über ein gemeinsames latentes Konstrukt bzw. mehrere Konstrukte erklärt werden können (vgl. Abbildung 5). Der Unterschied zur Faktorenanalyse besteht darin, dass für alle Items die gleiche Ladung vorausgesetzt wird (d. h. alle Items haben die gleiche Trennschärfe). Im einfachsten Fall nimmt man an, dass ein einzelnes latentes Konstrukt – etwa musikalische Kompetenz – die Ergebnisse erklären kann (vgl. Abbildung 5).

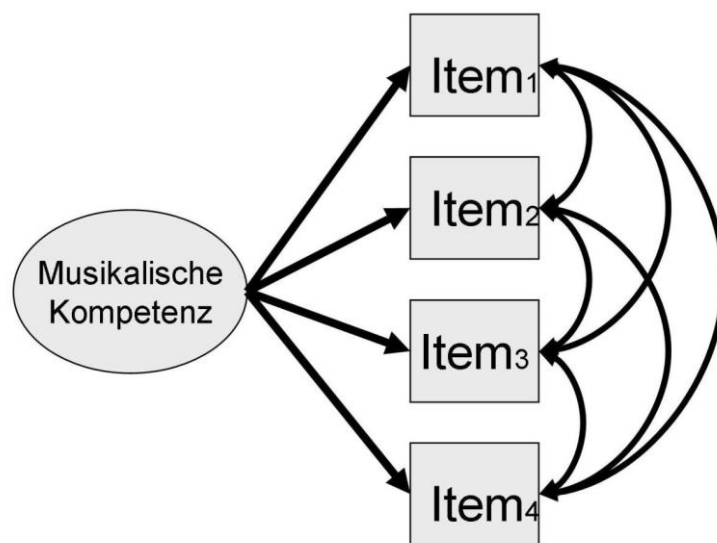


Abb. 5: Darstellung der latenten Variablen „Musikalische Kompetenz“, die hier durch vier untereinander korrelierende Items illustriert wird

Im Folgenden soll zunächst ein theoriegeleitetes, also konfirmatorisches Vorgehen vorgestellt werden. Dabei werden verschiedene theoretische Annahmen einander gegenübergestellt und anhand statistischer Kriterien verglichen, was einen Vergleich zwischen eindimensionalen und z. B. zwei-, vier-, fünf- oder sechsdimensionalen Modellen bedeutet. Die Items werden dazu *a priori* den Dimensionen zugeordnet. In einem weiteren Modell soll getestet werden, ob es Items gibt, die nicht eindeutig einer Dimension zugeordnet werden können, sondern z. B. neben der Wahrnehmungsfähigkeit auch Notations- und Fachterminologiekenntnisse erfordern (*between/within-Modelle*, u. a. Adams, Wilson, Wang 1997; Hartig, Höhler 2008).

Um entscheiden zu können, welches Modell die Struktur der Daten am besten abbildet, werden statistische Modellgütekriterien herangezogen. Allerdings sind Modellprüfungen immer relativ, denn gesucht wird nicht nach ei-

nem allgemeingültigen, überdauernden Modell, sondern nach einer Entscheidung zwischen alternativen Modellen hinsichtlich einer optimalen Anpassung an die Daten. Ein zusammenfassendes Modellgütekriterium bildet der so genannte CAIC-Index⁵ (Bozdogan 1987). Dieser berücksichtigt neben der Parameteranzahl und der Final Deviance (Likelihood)⁶ auch die Stichprobengröße. Das Modell mit dem kleinsten CAIC-Wert bildet die Daten am besten ab. Im Anschluss an die Identifizierung der Modellstruktur wird zur Überprüfung der Stabilität die Reliabilität der einzelnen Dimensionen bestimmt. Die Reliabilität (hier *EAP/PV Reliabilität*) gibt an, inwieweit eine Gruppe von Test-Items als Messung einer latenten Variablen betrachtet werden kann.

Zusammenhänge zwischen den Dimensionen

Nach Aufstellung der Modellstruktur sollen die Zusammenhänge zwischen den gefundenen Dimensionen genauer beleuchtet werden. Korrelationsanalysen sollen zeigen, ob die Zusammenhänge zwischen den Dimensionen so hoch sind, dass sie eine eindimensionale Modellstruktur nahe legen, oder ob die Dimensionen nur gering zusammenhängen und daher einzeln in einem mehrdimensionalen Modell angemessener abgebildet werden können. Im Falle einer Eindimensionalität wäre die Gesamtkompetenz auf einem singulären Kontinuum von niedriger bis hoher Kompetenz angeordnet. Bei einem mehrdimensionalen Modell würden die einzelnen Kompetenzdimensionen durch jeweils eine separate latente Dimension modelliert.

5 Ergebnisse

Modellierung der Kompetenz „Musik wahrnehmen und kontextualisieren“: Überprüfung der Modellgültigkeit und der Modellstruktur

Um die Modellgültigkeit mithilfe des Likelihood-Quotienten-Tests (Andersen-Test) überprüfen zu können, wurde die Stichprobe in zwei Untergruppen geteilt. Als Teilungskriterium wurde hier beispielhaft die Variable „Spielst du zur Zeit oder hast du früher ein Instrument gespielt?“ eingesetzt. Es wurde untersucht, ob die Abweichungen in den Itemantworten von Schülern mit und

5 Consistent Akaike's Information Criterion. $CAIC = -2 * \ln \text{Likelihood} + (\ln \text{Stichprobengröße} + 1) * \text{Anzahl Parameter}$.

6 Die Deviance ist ein Wert welcher angibt, wie gut das Modell die Daten abbildet.

ohne Instrumentalerfahrung so stark voneinander abweichen, dass die Modellgültigkeit in Frage gestellt werden muss. Zur Überprüfung dieser und weiterer Fragestellungen wurde die Software *ConQuest* (Wu et al. 2007) eingesetzt.

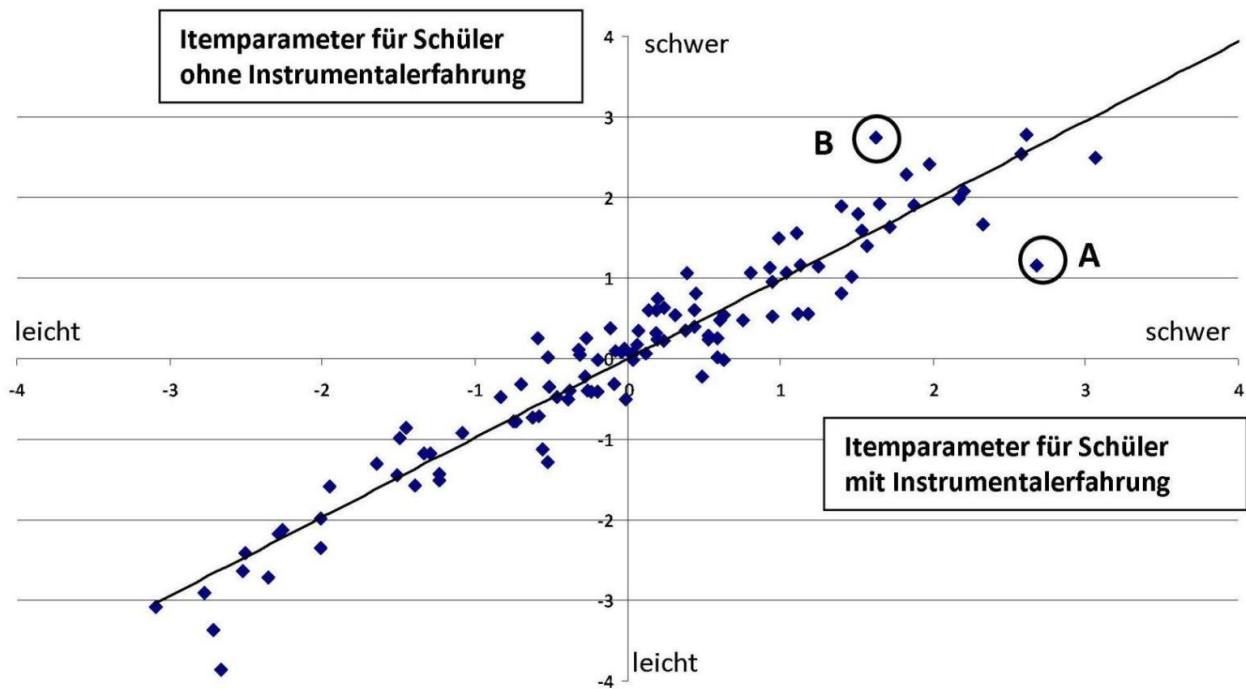


Abb. 6: Streudiagramm der Itemparameter gruppiert nach Schülern mit und ohne Instrumentalerfahrungen.

Legende

- A: Dieses Item ist für Schüler mit Instrumentalerfahrung schwerer (2,68) als für Schüler ohne Instrumentalerfahrung (1,16).
- B: Dieses Item ist für Schüler mit Instrumentalerfahrung leichter (1,63) als für Schüler ohne Instrumentalerfahrung (2,74).

Abbildung 6 verdeutlicht, dass die Itemparameter einzelner Items nicht auf der Geraden liegen und es somit Unterschiede zwischen den Gruppen gibt. Manche Items sind für Schüler mit Instrumentalerfahrung einfacher, was nicht sehr überraschend ist (s. Erklärung Abb. 6). Es muss aber überprüft werden, ob diese Unterschiede zu einer statistischen Verschlechterung des Modells führen. Tatsächlich mussten einige Items aus dem Finaltestheft für weitere Berechnungen ausgeschlossen werden. Außerdem mussten einige Items aufgrund des Teilungskriteriums „Geschlecht“ aus den weiteren Berechnungen gelöscht

werden. Nach Entfernung der problematischen Items kann von einer Modellgültigkeit ausgegangen werden.

Im Folgenden soll die Modellstruktur genauer betrachtet werden. Es wird überprüft, ob alle Items auf *einer* latenten Variable abgebildet werden können (vgl. Abbildung 5), oder ob sich die Items auf mehrere Dimensionen verteilen. Nach theoretischen Überlegungen werden, in Anlehnung an das erstellte Kompetenzmodell (s. Abb. 3), verschiedene Modellstrukturen gegenübergestellt. Bei dem 2-dimensionalen Modell werden der ersten Dimension alle Items zugeordnet, die den Bereich „Wahrnehmung und musikalisches Gedächtnis“ umfassen (vgl. Dimension 1 in Abbildung 3). Zu der zweiten Dimension gehören dann alle restlichen Items des Bereichs „Anwendung musikalischen Wissens auf Basis der Wahrnehmung“. Eine ähnliche Struktur suggeriert das abgebildete Kompetenzmodell (s. Abb. 3). Die 4-dimensionale Struktur ergibt sich aus den vier intendierten Dimensionen des Kompetenzmodells. Die kritische Bewertung von Musik ist in dem bisherigen Kompetenzmodell noch keiner Dimension zugeordnet. Demnach ist es nahe liegend, die Items, die sich mit der kritischen Bewertung von Musik beschäftigen, einer eigenen und somit der fünften Dimension zuzuordnen. Wie bereits beschrieben besteht darüber hinaus die Möglichkeit, Items mehreren Dimension gleichzeitig zuzuordnen (Within-Item-Modelle). Jedes Item wird somit daraufhin überprüft, ob es mehreren Dimensionen zugeordnet werden kann. Die Basis bildet das 4-dimensionale Modell, ein Item kann somit maximal vier Dimensionen gleichzeitig zugeordnet werden.

Mit der Software *ConQuest* wurden die Modellparameter (z. B. Final Deviance, Anzahl der Parameter) für die verschiedenen mehrdimensionalen Modelle berechnet. Aus der Stichprobengröße, der Anzahl der Parameter und der Likelihood (=Final Deviance) ergibt sich der CAIC-Index (Erläuterung der Formel s. o.). Somit ist das 4-dimensionale Modell im Vergleich mit dem 1-, 2-, 5-dimensionalen bzw. dem Within-Item-Modell das Modell mit dem kleinsten CAIC-Wert und erklärt die Daten am besten. Der Bereich der kritischen Bewertung von Musik bildet also keine eigene Dimension und muss noch einmal genauer betrachtet werden.

	1- dimensionales Modell	2- dimensionales Modell	4- dimensionales Modell	5- dimensionales Modell	4- dimensionales Within-Item- Modell
N	508	508	508	508	508
Anzahl Pa- rameter	102	104	111	116	111
Final Devi- ance	32229,20	32183,20	32044,76	32027,78	32126,01
CAIC- Index ⁷	32966,71	32935,17	32847,34	32866,52	32928,60

Tabelle 1: Modellvergleich zwischen einem 1-, 2-, 4- und 5-dimensionalen sowie einem Within-Item Modell.

Betrachtet man die Stabilität der vier Dimensionen, können folgende Reliabilitäten berichtet werden (Erwartungswert 1): $\alpha_{\text{Dimension 1}} = .81$, $\alpha_{\text{Dimension 2}} = .82$, $\alpha_{\text{Dimension 3}} = .78$ und $\alpha_{\text{Dimension 4}} = .70$. Anschließend wurden die Zusammenhänge zwischen den vier Dimensionen des Modells untersucht, um erneut die Sinnhaftigkeit eines 4-dimensionalen Modells zu überprüfen.

Zusammenhänge zwischen den Dimensionen

Wie in Tabelle 2 erkennbar, ergab sich der höchste Zusammenhang zwischen dem Bereich „Wahrnehmung und musikalisches Gedächtnis“ und der „Verbalisierung auf Basis der Wahrnehmung“. Auch zwischen der Dimension „Verbalisierung/Fachterminologie“ und „Notation“ (jeweils auf Basis der Wahrnehmung) zeigte sich ein hoher Zusammenhang, d. h. Schüler, die eine hohe Fähigkeit im Bereich der „Verbalisierung/Fachterminologie“ bewiesen, zeigten ebenfalls eine hohe Fähigkeit im Bereich „Notation“. Der geringste Zusammenhang bestand zwischen der Dimension „Notation“ und der Dimension „Historisch-kultureller Kontext“. Trotz einiger relativ hohen Korrelationen (s.

⁷ Das Modell mit dem kleinsten CAIC-Index bildet die Daten am besten ab.

Tab. 2), besonders zwischen Dimension 1 und 2, wurde inhaltlich und aufgrund des CAIC-Index die 4-dimensionale Struktur favorisiert.

	Dimension 1 (Reine Wahrnehmung)	Dimension 2 (Verbalisierung/ Fachterminologie in Verbindung mit Wahrnehmung)	Dimension 3 (Notation in Verbindung mit Wahrnehmung)	Dimension 4 (Historischer und kultureller Kontext in Verbindung mit Wahrnehmung)
Dimension 1	-			
Dimension 2	0.83	-		
Dimension 3	0.80	0.84	-	
Dimension 4	0.71	0.67	0.56	-

Tabelle 2: Latente messfehlerbereinigte Korrelationen zwischen den vier Dimensionen

6 Diskussion

Die Überprüfung der Modellgültigkeit ergab, dass nicht alle Items in das Raschmodell passten, d.h. dass sie die Anforderungen nicht erfüllten. Damit wird die Wichtigkeit des zweistufigen Pilotierungsdesigns (s. Abb. 4) sowie der sorgfältigen Analyse und Entwicklung der Items bestätigt.

Hinsichtlich der Modellstruktur konnte eine 4-dimensionale Struktur nachgewiesen werden, wobei die einzelnen Dimensionen hinreichend hohe Reliabilitäten besaßen. Die interne Konsistenz der vier Dimensionen war damit belegt. Lediglich die Reliabilität der vierten Dimensionen fällt mit $\alpha = .70$ vergleichsweise niedrig aus. Aufgaben dieser Dimension können somit nicht so deutlich trennscharf zwischen Schülern mit hoher und niedriger Fähigkeit unterscheiden.

Bei einem mehrdimensionalen Modell muss geklärt werden, wie hoch die einzelnen Dimensionen zusammenhängen, damit tatsächlich eine Eindimensionalität ausgeschlossen werden kann. Tabelle 2 zeigt, dass die höchste Korre-

lation ($r = 0.84$) zwischen den Dimensionen 2 (Verbalisierung/Fachterminologie) und Dimension 3 (Notation) zu finden ist. Deshalb ist es sehr wahrscheinlich, dass ein Schüler mit einer hohen Fähigkeit in Dimension 2 ebenfalls über eine hohe Fähigkeit in Dimension 3 verfügt. Zwischen den Dimensionen 3 und 4 zeigt sich die geringste Korrelation ($r = 0.56$), was aufgrund der völlig unterschiedlichen inhaltlichen Ausrichtung auch plausibel erscheint: Während es in Dimension 3 darum ging, Aufgaben im Bereich Notation auf Basis der Wahrnehmung zu lösen (z. B. ein gehörtes Musikstück einem Ausschnitt einer Partitur zuordnen), sollten die Schüler bei Aufgaben der Dimension 4 historische und kulturelle Aspekte der Musik berücksichtigen (z. B. ein gehörtes Musikstück dem Herkunftsland zuordnen und erläutern, warum diese Entscheidung getroffen wurde). Ein Vergleich unserer Ergebnisse mit denen anderer Studien (z. B. Bos, Lankes 2003) bestätigt die Entscheidung für Mehrdimensionalität. Bei einer Korrelation ab ca. $r = 0.90$ zwischen zwei Dimensionen müsste hingegen geprüft werden, ob es sich wirklich um ein mehrdimensionales Konstrukt handelt oder ob die beiden Dimension nicht vielmehr eine einzige latente Variable abbilden (vgl. Wu, Adams 2006).

Ein weiterer Vorteil des mehrdimensionalen gegenüber einem eindimensionalen Modell ist die Möglichkeit differenzierter Rückmeldungen der Ergebnisse an die Lehrkräfte. So kann eine Klasse beispielsweise im Bereich der Verbindung von Wahrnehmung mit Notation über sehr hohe und im Bereich der Verbindung von Wahrnehmung mit Verbalisierung/Fachterminologie hingegen nur über sehr geringe Fähigkeiten verfügen. Diese differenzierte Rückmeldung können die Lehrkräfte in ihre weitere Unterrichtsplanung einfließen lassen. Das eindimensionale Modell würde nur die Rückmeldung einer Gesamtkompetenz „Musik wahrnehmen und kontextualisieren“ mit entsprechend weniger diagnostischen und didaktischen Perspektiven ermöglichen.

7 Ausblick

In diesem Beitrag wurde die Überprüfung der Testgütekriterien nur am Rande erwähnt (z. B. Moosbrugger 2007, S. 8 ff.). Noch existieren z. B. keine standardisierten Tests in diesem Bereich, anhand derer die Validität des KoMus-Tests gemessen werden könnte. Der KoMus-Test wurde aber auf Basis eines theoretischen Modells erstellt, welches wiederum auf der Grundlage umfangreicher Curriculaanalysen (s. Knigge, Lehmann-Wermser 2008) entstanden ist. Dies stellt bereits ein Validitätskriterium dar, nämlich das der „Inhalts-“ bzw. „curricularen Validität“. Um die curriculare Validität möglichst objektiv abzu-

sichern, haben die Lehrkräfte der teilnehmenden Klassen im Rahmen der Pilotierungsstudie die Testaufgaben hinsichtlich Schwierigkeit, Bekanntheitsgrad, Relevanz und Lehrplanpassung beurteilt.

Wie eingangs beschrieben, gibt es neben dem vorgestellten Kompetenzstrukturmodell ein so genanntes Kompetenzniveaumodell. Hier werden für jede Dimension Kompetenzniveaus bestimmt. Dabei ist es besonders wichtig, wo die Übergänge von einem zum nächsten Niveau festgelegt werden; also ab welcher Fähigkeit ein Schüler das nächsthöhere Niveau erreicht hat. Die festgelegten Niveaus werden anschließend mit inhaltlichen Kriterien beschrieben, sodass deutlich wird, über welche Kompetenzen Schüler auf einem bestimmten Niveau verfügen, nicht aber auf dem nächst niedrigeren (vgl. Klieme et al. 2003, S. 24, Jordan (i.Vorb.)).

Bereits in den Modellgültigkeitsüberprüfungen (s. Abschnitt 5) wurde die Relevanz der Instrumentalerfahrung der Schüler deutlich. In weiteren Analysen sollte der Einfluss dieser und weiterer Hintergrundvariablen auf die Kompetenz im Bereich „Musik wahrnehmen und kontextualisieren“ differenziert untersucht werden. Neben den Instrumentalerfahrungen der Schüler könnten u. a. das Musikinteresse der Familie, die Anzahl der Jahre an Musikunterricht in der Schule und allgemeine Hintergrundvariablen wie der sozio-ökonomische Hintergrund oder das Geschlecht einbezogen werden.

Literatur

- Adams, R. J.; Wilson, M.; Wang, W.-C. (1997): The Multidimensional Random Coefficients Multinomial Logit Model. In: Applied Psychological Measurement, Jg. 21, H. 1, S. 1-23.
- Andersen, E. B. (1973): A goodness of fit test for the Rasch model. In: Psychometrika, Jg. 38, S. 123-140.
- Bähr, J. (2001): Zur Entwicklung musikalischer Fähigkeiten von Zehn- bis Zwölfjährigen. Evaluation eines Modellversuchs zur Kooperation von Schule und Musikschule. Göttingen: Cuvillier.
- Bähr, J. (2004): Bildungsstandards für den Musikunterricht? In: Ansohn, Meinhard; Terhag, Jürgen (Hg.): Musikkulturen - fremd und vertraut. Olfershausen: Lugert (Musikunterricht heute), S. 404-419.

- Bos, W.; Lankes, E.-M. (2003): Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich. Münster: Waxmann.
- Bozdogan, H. (1987): Model selection an Akaikes information criterion (AIC): The general theory and its analytical extension. In: Psychometrika, Jg. 52, H. 3, S. 345-370.
- Bruhn, H. (2005): Wissen und Gedächtnis. In: Oerter, R./Stoffer, T. H. (Hrsg.): Allgemeine Musikpsychologie. Göttingen: Hogrefe (Enzyklopädie der Psychologie, Serie VII, Bd. 1), S. 537-590.
- Bühner, M. (2008): Einführung in die Test- und Fragebogenkonstruktion. 2., aktualisierte und erw. Aufl., [Nachdr.]. München: Pearson Studium (psMethoden/Diagnostik).
- Hambleton, R. K.; Swaminathan, H. (2000): Item response theory. Principles and applications. 10. printing. Boston: Kluwer-Nijhoff (Evaluation in education and human services series).
- Hartig, J.; Höhler, J. (2008): Representation of Competencies in Multidimensional IRT Models with Within-Item and Between-Item Multidimensionality. In: Zeitschrift für Psychologie, Jg. 216, H. 2, S. 89-101. Online verfügbar unter <http://psycontent.metapress.com/content/w264233j2t034874/fulltext.pdf>, zuletzt geprüft am 10.09.2008.
- Hartig, J.; Klieme, E. (2006): Kompetenz und Kompetenzdiagnostik. In: Schweizer, K. (Hg.): Leistung und Leistungsdiagnostik. Mit 18 Tabellen. Heidelberg: Springer Medizin , S. 127-143.
- Hartig, J.; Klieme, E. (Hg.) (2007): Möglichkeiten und Voraussetzungen technologiebasierter Kompetenzdiagnostik. Eine Expertise im Auftrag des Bundesministeriums für Bildung und Forschung. Berlin: BMBF (Bildungsforschung, 20).
- Jank, W. (2001): Ist Musikhören wie Sprechenlernen? Musikalische Grundkompetenzen: Die Musikdidaktik muss von der Lerntheorie lernen. In: Musik & Bildung, H. 3, S. 31-39.
- Jordan, A.-K.; Knigge, J. (i. Druck): The development of competency models: An IRT-based approach to competency assessment in general music education. In: Brophy, T. S. (Hg.): The Practice of Assessment in Music Education: Frameworks, Models, and Designs. Chicago: GIA .

- Jordan, A.-K. (i. Vorb.): Empirische Validierung eines Kompetenzmodells für das Fach Musik – Teilkompetenz „Wahrnehmen und Kontextualisieren von Musik“ (Arbeitstitel).
- Kaiser, H. J. (2006): Bildungsoffensive Musikunterricht? Das Grundsatzpapier der Konrad-Adenauer-Stiftung in der Diskussion. Regensburg: ConBrio.
- Klieme, E.; Avenarius, H.; Blum, W.; Döbrich, P.; Gruber, H.; Prenzel, M. et al. (2003): Zur Entwicklung nationaler Bildungsstandards. Eine Expertise. Herausgegeben von Bundesministerium für Bildung und Forschung. (Bildungsforschung, 1). Online www.bmbf.de/pub/zur_entwicklung_nationaler_bildungsstandards.pdf, zuletzt geprüft am 10.07.2008.
- Knigge, J.; Lehmann-Wermser, A. (2008): Bildungsstandards für das Fach Musik: Eine Zwischenbilanz. In: Zeitschrift für Kritische Musikpädagogik, Sonderedition: Bildungsstandards und Kompetenzmodelle für das Fach Musik?, S. 60–98. Online verfügbar unter www.zfkm.org/sonder08-knigge-lehmannwermser.pdf, zuletzt geprüft am 27.08.2009.
- Knigge, J.; Lehmann-Wermser, A. (2009): Kompetenzorientierung im Musikunterricht. In: Musik und Unterricht, H. 94, S. 56–60.
- Knigge, J.; Niessen, A.; Jordan, A.-K. (2010): Erfassung der Kompetenz „Musik wahrnehmen und kontextualisieren“ mit Hilfe von Testaufgaben Aufgabenentwicklung und -analyse im Projekt KoMus In: Knolle, N. (Hg.): Evaluationsforschung in der Musikpädagogik. Essen: Die Blaue Eule (Musikpädagogische Forschung, 31).
- La Motte-Haber, H. de (2005): Modelle der musikalischen Wahrnehmung. Psychophysik - Gestalt - Invarianten - Mustererkennen - Neuronale Netze - Sprachmetapher. In: La Motte-Haber, H. de; Rötter, G. (Hg.): Musikpsychologie. Laaber: Laaber (Handbuch der Systematischen Musikwissenschaft, 3), S. 55–73.
- Lange, E. B. (2005): Musikpsychologische Forschung im Kontext allgemeinspsychologischer Gedächtnismodelle. In: La Motte-Haber, H. de/Rötter, G. (Hrsg.): Musikpsychologie. Laaber: Laaber (Handbuch der Systematischen Musikwissenschaft), S. 74–100.
- Lohmann, W. (1982): Ansätze zu einer objektiven Bewertung von Leistungen im Musikunterricht. Wolfenbüttel: Möser.

- Moosbrugger, H.; Kelava, A. (Hg.) (2007): Testtheorie und Fragebogenkonstruktion. Mit 43 Tabellen. Heidelberg: Springer Medizin (Springer-Lehrbuch).
- Niessen, A.; Lehmann-Wermser, A.; Knigge, J.; Lehmann, A. C. (2008): Entwurf eines Kompetenzmodells 'Musik wahrnehmen und kontextualisieren'. In: Zeitschrift für Kritische Musikpädagogik, Sonderedition: Bildungsstandards und Kompetenzmodelle für das Fach Musik?, S. 3–33. Online verfügbar unter <http://www.zfkm.org/sonder08-niessenetal.pdf>, zuletzt geprüft am 20.03.2010.
- Richter, C. (2007): Bildungsstandards und Kompetenzformulierungen im Fach Musik. In: Labudde, Peter (Hg.): Bildungsstandards am Gymnasium. Korsett oder Katalysator? Bern: h.e.p. , S. 273–281.
- Rost, J. (2004): Lehrbuch Testtheorie - Testkonstruktion. 2., vollst. überarb. und erw. Aufl. Bern: Huber (Psychologie Lehrbuch).
- Serafine, M. L. (1988): Music as cognition. The development of thought in sound. New York: Columbia University Press.
- Stoffer, T. H. (2005): Aufmerksamkeitsprozesse beim Musikhören: Wissensunabhängige und wissensabhängige Selektionsprozesse. In: Oerter, R./Stoffer, T. H. (Hrsg.): Allgemeine Musikpsychologie. Göttingen: Hogrefe (Enzyklopädie der Psychologie, Serie VII, Bd. 1), S. 591–656.
- van der Linden, W. J.; Hambleton, R. K. (1997): Handbook of modern item response theory. New York, NY: Springer.
- Weinert, Franz E. (Hg.) (2001): Leistungsmessungen in Schulen. Weinheim: Beltz.
- Wu, M.; Adams, R. J. (2006): Modelling Mathematics Problem Solving Item Responses Using a Multidimensional IRT Model. In: Mathematics Education Research Journal, Jg. 18, H. 2.
- Wu, M.; Adams, R. J.; Wilson, M.; Haldane, S. A. (2007): ConQuest - Generalised item response modelling software. Camberwell: Australian Council for Educational Research.

Metaphorischer Sprachgebrauch im Unterricht

Überlegungen zur Evaluierung der Schülersprache

„Für den Transfer von einer Sprachebene in die andere ist die Metapher unverzichtbar. Darüber ist in der Musikpädagogik nur sehr wenig nachgedacht worden. Der Gebrauch von Metaphern erfordert eine Sprachphantasie, die nicht einmal allen Musiklehrern hinreichend verfügbar ist. Sie ist nicht mit einer unverbindlichen, ‚blumigen Rede‘ zu verwechseln. Metaphern sind bildhaft-vorsymbolische Brücken der Verständigung und verweisen in den Raum des Unsagbaren, der für die Künste mindestens ebenso wichtig wie das (spärlich) Sagbare ist. Zugleich tragen sie vieles mit, was unsere Lebenswelt ausmacht.“¹

Karl-Heinrich Ehrenforth gestattet sich diesen Seitenblick auf das Metaphorische im Rahmen von Überlegungen, die die Erträge einer lebensweltlichen Orientierung des Musikunterrichts in den Kontext einer pädagogischen Praxis stellen. Was in diesem Zusammenhang quasi en passant über das Metaphorische gesagt wird, soll nun in einen größeren theoretischen Rahmen gestellt und dabei gleichzeitig an Hand einer konkreten Unterrichtssituation hinsichtlich der Praxisrelevanz geprüft werden. Die Metapher ist für Ehrenforth ein Transportmittel der Verständigung, das getrennte Sphären überbrückt: Dienlich ist metaphorische Sprache, um einen Ausgleich zwischen sinnlichem Erleben und kognitivem Verstehen zu erzielen. Ihren Mehrwert entfaltet sie dabei zugleich im (interpretativen) Erschließen des (begrifflich) Unsagbaren, als auch in der Annäherung an den Gegenstand aus der Perspektive der Schüler. Im letzteren Fall besitzt sie den Status der Vorläufigkeit und hilft dort aus, wo die etablierten Begriffe fehlen. Implizit mitgedacht ist hier wohl, dass die angesprochene Sprachphantasie nicht allein das Bilden, sondern auch das Verstehen der Metapher betrifft, das ebenso sehr von schöpferischem Bemühen geleitet ist wie das Hervorbringen. Insgesamt kommt dabei dem Metaphori-

1 Ehrenforth 2001, S. 55.

schen eine vornehmlich vermittelnde Position zu: Lebensweltlich gefärbt vermittelt sie zwischen individuellem Erleben und strukturierter Ordnung.

In den folgenden Darlegungen soll der Versuch unternommen werden, Ehrenforths Anregung aufzunehmen und über die Bedeutung der Metapher für das Musikverstehen nachzudenken.

Dabei soll

- der Status der Metapher genauer beleuchtet werden: Ihr Weg aus dem Reservat von Rhetorik und Dichtung hin zum Nährboden des Denkens. Der Wandel von einer uneigentlichen Verkleidung wörtlicher Bedeutung zum sprachlich feststellbaren Niederschlag einer kognitiven Erscheinung.
- die Metapher als ‚Schaltstelle‘ im pädagogischen Kontext betrachtet werden, um das Konzept der kognitiven Metapher genauer zu umreißen und vor dem Hintergrund eines Metaphernverständnisses abzugrenzen, das in der Metapher eher ein Medium der Wissensübertragung, allenfalls einen Katalysator des Verstehens sieht. Dies soll schlagpunktartig, durch die Gegenüberstellung zweier Positionen von Walter Herzog und Peter Gansen erfolgen, die im Abstand von 25 Jahren eine Standortbestimmung versuchen und Einblicke in verschiedene Aspekte des Metaphorischen im pädagogischen Kontext gewähren.
- eine Strategie entwickelt werden, die hilft, das Denken in Metaphern methodisch zu kontrollieren, planbar und (ver)mittelbar zu gestalten.
- die Verschaltung der kognitiven Metaphern im Lernprozess anhand einer Unterrichtssituation genauer beleuchtet werden. Im Mittelpunkt stehen hier Verstehensprozesse: Strategien, die Wissen, das in einer intuitiv verwendeten Metapher verborgen ist, zum Vorschein bringen.

Grundsätzliches über das Metaphorische

Gestützt durch die zweieinhalb Jahrtausende alte Lehrtradition der Rhetorik ist die Metapher eine Erscheinung des Sprachgebrauchs mit Resonanzen im Denken: Die Metapher ist „vom Dichter geprägt.“² Diesen effektvollen Zierrat in poetischen Texten gesteht man denen zu, die es mit der Wahrheit nicht so genau nehmen müssen: Sprachbilder haben ihren angestammten Ort in der Dich-

2 Aristoteles 2006, 1457 b, S. 69.

tung, auch der Redner nutzt affektive Wirkungen dieses Sprechens, die „täuschende, uneigentliche Verkleidung des Wörtlichen“³, um seine Adressaten zu überzeugen oder in unseriöser Weise suggestiv zu beeinflussen. In der Wissenschaft hingegen, auch in der sprachlichen Auseinandersetzung mit den schönen Künsten, wird die Sprache als neutrales Medium der Informationsvermittlung gesehen. In Abgrenzung von „eigentlicher“ (also begrifflicher) und „uneigentlicher“ (d.h. metaphorischer) Rede gilt die Metapher aufgrund ihrer schwer zu fassenden Vieldeutigkeit als eine Anomalie, die zu einer zu bekämpfenden Krankheit ausufern kann. Metaphern gelten hier immer nur als zweitbeste Lösung, die sich durch wörtliche und damit wahrheitsfähige Begriffe ersetzen, d.h. substituieren lassen. Fazit: Die Metapher gehört in die Literatur, die Sprache der Wissenschaft erfordert den akademisch gereinigten Begriff. Harald Weinrich formuliert es etwas freundlicher: „Eine gewisse Enthaltsamkeit in der Metaphorik“ gehört „zum stilistischen Ethos vieler Wissenschaftler.“⁴

Das Konzept der Substitution oder Übertragung vermag das kreative Potential der Metapher nicht zu erfassen. In Opposition zum aristotelischen Substitutionsgedanken favorisieren heutige Diskussionen der Metapher ein dynamisches Verständnis der Bedeutungsentstehung, bei der verschiedene semantische Felder miteinander verknüpft werden. Harald Weinrich hat dafür die Begriffe Bildspender und Bildempfänger eingeführt.⁵ Beide werden durch eine Ansammlung von Assoziationen miteinander verbunden: Mit der Metapher geht eine Wahrnehmungslenkung einher, die bestimmte Eigenschaften verstärkt, während andere vernachlässigt werden. Der Sinn des metaphorischen Sprechens ist nicht, stabile Bedeutungen zu schaffen. Bedeutung entsteht vielmehr in einem dynamischen Prozess der Wechselwirkung, die neue Horizonte eröffnet und Bedeutung stiftet.

Die derzeitige Konjunktur des Metaphorischen gründet sich auf Lakoff und Johnson, die die Metapher als einen integralen Bestandteil unseres alltäglichen Sprechens und Denkens herausstellen. Sie kehren somit die altbewährte Ansicht auf das Metaphorische um: „Die Metapher ist primär eine Sache des Denkens und Handelns und erst sekundär eine sprachliche Angelegenheit.“⁶

3 Debatin 2005, S. 32.

4 Weinrich 1976, S. 188.

5 Weinrich 1976.

6 Lakoff & Johnson 2004, S. 177.

Die Metapher gilt nicht mehr als eine Erscheinung des Sprachgebrauchs mit Resonanzen im Denken, sondern als eine Erscheinung des Denkens mit Resonanzen im Sprachgebrauch. Die sprachlich feststellbare Erscheinung ist also lediglich der sichtbare Niederschlag kognitiver Konzeptbildung, die Metapher wird so zu einer „anthropologischen Universalie“⁷:

„Stand vordem der Beitrag der Metapher zum Stil im Vordergrund, so wird aus dieser Perspektive heraus ihr Beitrag zu unseren kognitiv-sprachlichen Fähigkeiten und unserer geistigen Kraft ersichtlich. Deutlich wird vor allem ihre Kraft, unser Denken, unsere Emotionen, unsere imaginativen Fähigkeiten und unsere Sprache produktiv interagieren zu lassen, sowie auch ihr Potenzial, unseren innersten Gefühlen und abenteuerlichsten Vorstellungen eine Struktur und einen kommunizierbaren Sinn zu verleihen.“⁸

Lakoff und Johnson versuchen zu zeigen, „daß die menschlichen Denkprozesse weitgehend metaphorisch ablaufen“⁹, entfalten so das schöpferische Potential der Metapher. Demnach können Sachvorstellungen verschiedener Bereiche aneinander gekoppelt werden und eine Erfahrung partiell in Begriffen einer anderen Erfahrung strukturiert werden. Die Metapher ist also nicht mehr ausschließlich ein Transportträger, sprachliches Medium, sondern „Werkzeug des Denkens“¹⁰.

Aus Bildern lernen: Die Metapher als vermittelnde Instanz zwischen Wissen und (Noch)-Nichtwissen

Während das kognitive Potential der Metapher in den Sprach- und Naturwissenschaften bereits allgegenwärtig ist, stellt die Pädagogik sie noch vornehmlich in den Dienst der Wissensvermittlung, um Laien Fachwissen nahezubringen. Die Metapher erscheint als „Dekor mit didaktischen Endzwecken“¹¹, um

7 Kohl 2007, S. 119.

8 Kohl 2007, S. 119f.

9 Lakoff & Johnson 2004, S. 14.

10 Drewer 2003.

11 Nieraad 1977, S. 10.

*„Dem, der nicht viel Verstand besitzt,
Die Wahrheit, durch ein Bild, zu sagen.“¹²*

Diese von Gellert in der Fabel von der Biene und der Henne vorgetragene Verknüpfung von Anschaulichkeit und Bildlichkeit fügt sich in den Lehrplan des Aristoteles, der in der Metapher gerade diese Wirkung des „Vor-Augen-Führens“ hervorhebt. Dieser Status der vermittelnden Instanz zwischen Wissen und Nochnichtwissen prägt bis heute den Umgang mit der Metapher in der Pädagogik:

„Für die Vermittlung sind Metaphern geeignet, Interesse zu wecken und einen Bezug zwischen Lernenden und Gegenstand herzustellen, denn wenn der Herkunftsbereich der Metapher vertraut ist, können Lernende auf Bekanntes zurückgreifen, um sich den Zielbereich zu erschließen.“¹³

Das hier ausgebreitete didaktische Potential der Metapher soll nun an Hand eines protokollierten Unterrichtsgesprächs einer 11. Klasse über die Eingangstakte der 1. Sinfonie Beethovens verdeutlicht werden:

„Ein kurviger Weg.“

Diese reflexartige Schüleräußerung dokumentiert zunächst Orientierungswissen. Eine erste tastende Äußerung, die zeigt, wie sich Schüler den Sachverhalt denken. Dabei bleibt zunächst offen, ob sie sich auf die einleitenden Bläserakkorde oder auf die Streichermelodik bezieht. Sie versucht ihren Eindruck als Ganzes in Worte zu fassen. Die Äußerung wirkt zunächst eher konventionell: Es entspricht ja durchaus einer gängigen Konzeptionalisierung, Musik in ihrem flüchtig, zeitlichen Verlauf im Raum anzuordnen. Die ‚Kurven‘ ließen sich in der Partitur nachzeichnen und veranschaulichen. Derartige Anschauungsäquivalente sind in der Umgangssprache legitim, in der Sprache der Wissenschaft haben sie aufgrund ihrer vermeintlichen Unschärfe und exponierten Subjektivität keinen Platz. Auch Ehrenforth bleibt zurückhaltend: Die Metapher erscheint ihm lediglich als „bildhaft-vorsymbolische Brücke der Verständigung.“ „Ein kurviger Weg“ ist in diesem Sinne eine Vorstufe der Begriffsbildung, eine Metapher, die dort aushilft, wo es den Schülern (noch) am Begriff fehlt. So ließe sich mit Hilfe sprachlicher Bilder ein direkter Zu-

12 Zit. bei Nietzsche 1988, S. 92.

13 Peyer 2006, S. 25, in Anlehnung an Drewer 2003, S. 92.

gang zur musikalischen Fachsprache eröffnen, der melodische und harmonische Verlauf beschreiben, der umgangssprachliche Volkston in die Sprache der Wissenschaft übersetzen. Dann spräche man nicht mehr von einer Kurve, sondern von Melodiebögen, auf- und absteigenden Melodieverläufen. Eine externe Sinnstiftung („ein kurviger Weg“) ist nun in eine interne Metaphorik (Melodiebogen) überführt, die den Schüler zu zentralen musikalischen Strukturmomenten führt.

An Hand dieser Schüleräußerung sind bisher Phänomene beschrieben worden, die Lakoff und Johnson als „isolierte Ähnlichkeiten“¹⁴ bezeichnen würden: Einzelne Übereinstimmungen zwischen Bildspender und Bildempfänger werden herausgegriffen und damit die Metapher auf der Ebene des analogen Vergleichs abgehandelt, letztlich also substitutionstheoretisch aufgefasst: Zugrunde liegt hier die Vorstellung der Substitution einer uneigentlichen (also metaphorischen) durch eine eigentliche (d.h. wörtliche) Bedeutung. Die Metapher bleibt ein Zwischenschritt, eine bebilderte Vorform zum eindeutigen, begrifflichen Denken. Die Expertensprache einer bereits erschlossenen Welt wird so an die Lebenswelt des wahrnehmenden Subjekts angebunden oder umgekehrt: eine metaphorische Beschreibung wird in die Begriffssprache übersetzt. So ließe sich das Schlusswort aus Gellerts Fabel umdrehen: Ist der Verstand erst einmal durch Wissen angereichert, macht dies das ursprüngliche Bild überflüssig. Alle Auseinandersetzung mit der Metapher hat zwar den Markt der pädagogischen Möglichkeiten bereichert, man schätzt die Metapher wegen ihrer Anschaulichkeit, letztlich bleibt die Metapher aber eine sprachliche Erscheinung, ein didaktisch-methodisches Mäntelchen, das man je nach Kenntnis oder Wissenstand der Schüler umlegen muss oder wieder ablegen kann. Reduziert auf diese Bedingungen der Begriffssprache verbleibt sie hier also letztlich auf einer Ebene der Analogie, auf der bereits angedeuteten didaktisch-illustrativen Ebene, hat somit lediglich Vermittlungsqualität. Auch die komprimierende Funktion einer Metapher, ihre Fähigkeit, die flüchtig verklingende Musik in eine fassliche Form zu bringen, gehört in diesen Zusammenhang: Die Erinnerungsprojektionen, die zwischen dem verflogenen Klang und dem reflexiven Sprechen darüber vermitteln. Dieser Vorgang der Analogiebildung lässt sich aus pädagogischer Sicht mit Werner Herzog zusammenfassen:

14 Lakoff & Johnson 2004, S. 177.

„Ein schulischer Stoff, der in einer begrifflichen Sprache darstellbar ist, lässt sich allmählich von seinen metaphorischen Krücken befreien und den Schülern ‚wörtlich‘ nahebringen.“¹⁵

Für Herzog steht demnach die vermittelnde Instanz zwischen Wissen und (Noch)-Nichtwissen im Vordergrund: Die Metapher erscheint als ein irrationales Hilfsmittel zur Anregung von rationalen (d. h. nichtmetaphorischen) Strukturen. Sie ist Ausdruck der Unvollkommenheit unseres Denkens im Bereich des Noch-nicht-Verfügbaren. Ihre Relevanz entwickelt die Metapher für Herzog eher auf anderem Gebiet, in der Entfaltung des Kreativen und Phantastischen,¹⁶ ein Terrain, das ohne Metapher nicht begehbar wäre. Auch Ehrenforth verortet das Potential der Metapher vornehmlich im „Raum des Unsagbaren“ und bringt sie dort ins Spiel, wo die begrifflichen Kompetenzen allgemein versagen: Im Raum des Ahnungsvollen, Nebulösen.

Das schöpferische Erkenntnispotential der Metapher im Grenzbereich von Wissen und Nichtwissen

Mit Blick auf die Forschungssituation stellt Peter Gansen fest, dass sich die Beschäftigung mit der Metapher in der Pädagogik eher in einer selbstreferentiellen Diskussion über pädagogische Grundbegriffe und deren inhärenten Basismetaphorik im eigenen Fach erschöpft: „Der zurückhaltende, um freies Wachstum besorgte Gärtner und der manipulierende, das Verhalten formende Lerntechnologe umschreiben bildhaft die beiden Pole des breiten Spektrums pädagogischer Haltungen.“¹⁷ Untersucht werden etwa die verborgenen Handlungsanweisungen eines um freies Wachsen und Reifen seiner Pflanzen besorgten Pädagogen als „Gärtner“, gerade wenn der seine Pflanzen nicht nur „wachsen“ lässt, sondern auch „beschneiden“ und „kranke Triebe“ entfernen muss: „Gerade in der Sprache der Pädagogik lassen sich metaphernanalytisch betrachtet viele begriffliche Ungenauigkeiten und Verschleierungen der Erziehungswirklichkeit bis hin zu ideologisch geprägter Gesinnungsrhetorik finden.“¹⁸ – Betrachten wir nun die Sprache der Musikwissenschaft, akzeptieren wir die metaphorischen Ursprünge der etablierten Begriffssprache: Musik ist

15 Herzog 1983, S. 313.

16 Vgl. hierzu Herzog 1983, S. 316ff.

17 Herzog 1983, S. 300.

18 Gansen 2009, S. 328.

Gewebe, Sprache, Energie, Struktur usw. Dabei lassen wir uns ebenfalls als Gefangene unserer Metaphern von dieser basalen Konzeptionalisierung führen und vergessen dabei, dass auch etablierte Routinen im Kern an ihre Modellfunktion und den damit fixierten Denkweisen gebunden bleiben. In diesem Zusammenhang ließe sich auch Ehrenforths Beschreibung der Bedeutung metaphorischen Sprechens für das Musikverstehen hinterfragen: Die „Brücke des Verstehens“ ist ein Bild, das Statisches suggeriert, ursprünglich getrennte Bereiche zusammenfügt, Lebensweltliches mit einem fest konstituierten Gegenstand einer objektiven Welt verbindet (aber nicht etwa paart oder vereinigt), den es zu erkennen, benennen und zu verstehen gilt.

Es fehlt insgesamt das Bestreben, die Diskussion über die Metapher im pädagogischen Kontext aus dieser Metaebene zu befreien und praxisrelevant werden zu lassen. Ausgehend von diesem weißen Fleck versucht Gansen nun, Perspektiven für notwendige Unterrichtsforschung aufzuzeigen:

„Wie erfolgt Aneignung von Wissen unter metaphorologischer Perspektive? [...]

(Wie) lässt sich metaphorisches Denken lehren, lernen und fördern? [...]

Welche Möglichkeiten bietet die Metaphorik von Sprache und Bildern, um ästhetische Lernprozesse zu initiieren und auszugestalten? “¹⁹

Wir haben die Schüleräußerung „ein kurviger Weg“ zunächst gelesen als eine didaktische, erklärende, wissensvermittelnde Metapher, die den Blick auf den Gegenstand dabei vorschnell determiniert:

„Durch das andauernde Thematisieren der didaktisch illustrativen Metaphern wird die Aufmerksamkeit auf die Oberflächenmetaphorik gerichtet, von der die ‚normale‘ Wissenschaftssprache sich als wörtlich abgrenzt, wodurch aber gleichzeitig die epistemisch-metaphysische Tiefenmetaphorik der Wissenschaftssprache verdeckt wird.“²⁰

19 Gansen 2009, S. 340.

20 Debatin 1995, S. 218f.

Natürlich beruht die Metapher partiell auf den bereits beschriebenen isolierten Ähnlichkeiten, sie können auch Auslöser für ein Bild sein. Zu wichtigsten Ähnlichkeiten gehören jedoch diejenigen, die durch die Metapher erst hergestellt werden: Nun geht es also darum, die Metapher in ihrer ganzen assoziativen Reichweite zu betrachten. Eingenommen wird dabei eine konstruktivistische Perspektive, die davon ausgeht, dass Metaphern zu einer subjektorientierten Auseinandersetzung mit gegebenen Realitäten herausfordern, diese umdeuten und umformen können: Die Metapher erscheint als ein kreatives Element, dienlich zur Artikulation eines neuen kognitiven Gehalts. Die Metapher als „Brücke der Verständigung“ (Ehrenforth) überbrückt nicht ausschließlich die Distanz zwischen unterschiedlichen Sprachebenen. Um das produktive Vermögen der Metapher zu ergründen, gilt es nun, die Metapher aus diesem Reservat zu ziehen und ins Zentrum der Beziehung zwischen Denken und Sprache zu führen. Sie löst einen Brückenschlag in unerforschte Gebiete aus, überschreitet dabei Grenzen des begrifflich Markierten und erzeugt Perspektiven im unerforschten, fremden Terrain. Ehrenforth spricht auch diesen Gesichtspunkt an und verweist auf den Ausblick auf das „Unsagbare“, das sich nur in der Sphäre des Metaphorischen ansprechen lässt und sonst unzugänglich bliebe, grenzt metaphorische Sprache von einer unverbindlichen ‚blumigen Rede‘ ab. Er impliziert damit eine Unterscheidung zwischen zugelassenen und unseriösen, guten und bösen Metaphern. Kann hier die Metapher, die mit ihrem ‚Als-ob-Status‘ und ihrer Lizenz zum ‚uneigentlichen‘ Sprechen keinen Anspruch auf einen rationalen Wahrheitsgehalt erhebt, den schlechten Ruf programmatischer Sinnzuschreibungen neu verhandeln? Erscheint die „blumige Rede“ in einer poetisierenden Paraphrase gar in einem anderen Licht, wenn man der Gefahr fester programmatischer Sinnzuschreibungen entgeht, indem man ihr den Freiraum einer Metapher gewährt?

Die Metapher zwischen Rationalität und intuitiv-vorgreifendem Wissen

Eine metaphorische Sprache lässt sich im Unterricht nicht verordnen: „Metaphern drängen sich eher auf, sie kommen.“²¹ Auch die kognitive Metapherntheorie krankt an einer Vagheit und Unbestimmtheit, was gerade zentrale Aspekte ihrer Theorie angeht, nämlich wie nun der Schritt von konkreten sprach-

21 Kurz 2004, S. 21.

lichen Metaphern zu einem Rückschluss auf den musikalischen Gegenstand und den zu Grunde liegenden Konzepten vollzogen wird.²²

Es gilt nun, den Geltungsbereich der Metapher auszuweiten, metaphorische Prozesse auch auf das elementare Sprechen des „(spärlich) Sagbaren“ (Ehrenforth) anzuwenden. Dabei ist zu zeigen, dass die Metapher nicht nur vermittelnd wirkt, sondern selbst auf allen Ebenen Erkenntnisvorgänge initiiert, sinn- und strukturstiftend zu wirken vermag, damit also dem Denken voraus ist. Es geht darum, dieses Denken einzuholen und jenes Wissen zum Vorschein zu bringen, das in der Metapher (noch) verborgen ist. Bernhard Debatin spricht hier vom „rationalen Vorgriff“ der Metapher:

„Zusammenfassend möchte ich diese Überlegungen in die These kleiden, daß die lebendige Metapher ein rationaler Vorgriff ist, daß sie lebensweltliche, oft vorbegriffliche Vorstellungen zum Ausdruck bringt und deren Sinn konstruiert, daß sie fiktive imaginäre und utopische Ideen entwirft und daß sie als heuristisches Mittel zur Artikulation von neuen kognitiven Gehalten dient.“²³

Hiermit verbindet Debatin zwei Seiten *einer* Medaille, die das Metaphorische stets begleiten: Die intuitiv-vorgreifenden Momente, also das kreative, sich dem gelenkten Zugriff entziehende Potential und das welterschließende Vermögen, als Ursprung begrifflichen Denkens. Seine Grundthese, die er in seiner Arbeit zur Rationalität der Metapher exemplifiziert, lautet mit Lichtenberg formuliert und auf eine prägnante Formel gebracht:

„Die Metapher ist viel klüger als ihr Verfasser und so sind es viele Dinge.“²⁴

Letztlich drückt dies aus, was Schenker als „den ahnenden Begriff des Totaleindrucks“ bezeichnet, „weil das Bewußtsein so klar, so tief und umfassend, als es nötig ist, überhaupt noch nicht hineinleuchten kann.“²⁵

Mit den Begriffen „Rationalität“ und „Vorgriff“ nimmt Debatin zwei Kennungen des Metaphorischen vor, die sich zunächst zu widersprechen scheinen:

22 Vgl. hierzu Baldauf 2000, S. 124f.

23 Debatin 1995, S. 137.

24 Lichtenberg 1968, F 369.

25 Schenker 1990, S. 99 u. S. 98.

Schließlich weisen vorgreifende Strukturen über die gesicherte Erkenntnis der Ratio hinaus. Für Debatin erhält sich die Metapher den Status einer „auf der Imagination beruhenden Rationalität.“²⁶ In diesem konstruktivistischen Prozess bilden Metaphern die Welt nicht ab, sondern bieten subjektive Bedeutungskonstruktionen auf der Basis eines offenen Reflexionsaktes. Gleichzeitig gilt es, diese vorgreifenden Sinnzuschreibungen in einem reflexiven Prozess der Metaphernkontrolle an den musikalischen Gegenstand anzubinden, die „vorgreifend-umherschweifende Funktion der Metapher methodisch zu kontrollieren und zu reflektieren.“²⁷ Sich abgrenzend und positionierend „gegenüber Traditionen, die in ihr eine irrationale und überflüssige Erscheinung sehen, und gegenüber Traditionen, die in ihr eine irrationale, obzwar notwendige Erscheinung erblicken, [indem er] mit Nachdruck darauf insistiert, daß sie eine notwendige und rationale Erscheinung ist.“²⁸ Durch die „reflexive Metaphorisierung“²⁹, die Ausnutzung der metaphorischen Leistung zu Zwecken der Metaphernreflexion, wird die Metapher nicht verbannt, vielmehr wird ihre Reichweite, „die metaphorische Kraft, sowie die Möglichkeit und das Potential alternativer metaphorischer Beschreibungen geprüft.“³⁰ Hierzu rechnet Debatin das Sichtbarmachen der Hintergrundmetaphorik zur Wiederbelebung „toter“ Metaphern, das Aus- und Weiterspinnen einer Metapher, aber auch die Grenzbestimmung ihrer Reichweite durch die Konstruktion von Bildbrüchen und paradoxer Metaphern. Auch die „blumige Rede“ kann im Rahmen einer solchen Metaphernkontrolle ihre Rationalität erweisen: Voraussetzung hierfür ist jedoch die Verschränkung von (vorgeblich) begrifflich-objektiver Analyse und (scheinbar) wuchernder subjektiver Metaphorik, die Verbindung zwischen intuitiv Vorgreifendem und rational einholbarem Wissen. Das Ergebnis ist eine klangsinnliche Beschreibung der Formung, die begrifflich-technische Bestimmung und poetische Fiktion verbindet.

26 Lakoff & Johnsson 2004, S. 269.

27 Debatin 1995, S. 156f.

28 Rolf 2005, S. 259.

29 Ebatin 1995, S. 165f. u. Debatin 2005, S. 39ff.

30 Debatin 2005, S. 41.

Der Metaphorische Prozess: Die initiale Metapher, ihre Verdichtung und Konsolidierung

Aufgehoben finden sich die Momente der Metaphernkontrolle im bereits angedeuteten Spannungsfeld zwischen Orientierung und Konsolidierung in einem Modell in Anlehnung an Anil K. Jain³¹, der in seiner Studie über die Metapher der sozialen Landschaft drei Phasen des metaphorischen Prozesses unterscheidet. Ein Versuch, amethodische Verstehensprozesse, Unterrichtssituationen, die sich aus einem zufälligen Kategorienfehler ergeben, zu kontrollieren: Zunächst bildet die initiale Metapher den Ausgangspunkt, den ersten orientierenden Zugang, der einem erlaubt, in unbekannte Welten hineinzublicken. Hier schließt sich eine Phase der assoziativen Ausgestaltung und Verdichtung an, um abschließend die weiter detaillierte und ausgemalte Metapher in kritischer Reflexion am ursprünglichen Gegenstand zu exemplifizieren.

Dieses Verfahren soll nun auf unsere initiale Metapher „ein kurviger Weg“ angewendet werden. Der Anbindungsprozess zwischen Metapher und dem musikalischen Gegenstand ergab sich bisher durch ein In-Beziehungsetzen einzelner Bildelemente mit der musikalischen Struktur. Als ursprünglicher und vorläufiger Ausdruck des ästhetischen Verstehens nimmt die Metapher jedoch nicht den Platz einer fehlenden Vokabel ein: Wenn sich ein Schüler in dieser Form äußert, tut er dies nicht, weil ihm der Begriff „Melodiebögen“ fehlt, er nimmt die Musik in ihrer Gesamtheit, also gestalthaft wahr. Das Einkleiden in die terminologische Begrifflichkeit blendet die ursprüngliche Wahrnehmungssituation aus, und geleitet von diesen Begriffen, lässt sie sich das Wahrgenommene nur schwer wieder einholen. Wir haben verloren, was wir vorher schon besessen haben. Maßgeblich ist hier eine Unterrichts Atmosphäre, die Metaphern nicht von vornherein als eine Sprache der zweiten Wahl degradiert. Voraussetzung ist hier (bei Schülern *und* Lehrern) ein Bewusstsein, bei dem Metaphern nicht nur geduldet, sondern ausdrücklich erwünscht sind und den Lernprozess in jeder Phase des Unterrichts begleiten.

Will man nun das Wissen, das in dieser Metapher wurzelt, zum Vorschein bringen, kann dies nur durch eine Wurzelbehandlung erfolgen. Daher gilt es nun, die ursprüngliche initiale Metapher in einem zweiten Schritt weiter auszumalen, um die von einer vordergründigen Übertragung verdrängten Aspekte auszuspielen. Die Schüler sind in dem nun weiter zu verfolgenden Unterrichts-

31 Jain 2002, S. 16ff.

gespräch „die Komplizen des metaphorischen Geschehens“, ihre „Rezeptionskompetenz und Kombinationsgabe vollendet den Effekt“³²:

*„In der Kurve weiß man nicht, wie es hinterher weitergeht.“
„In einer Kurve muss man bremsen.“*

Neben der bereits zugestandenen komprimierenden Funktion der Metapher geht es hier also darum, den Blick auf den bildspendenden Bereich zu richten und gleichsam eingefahrene Sichtweisen zu entgrenzen. Denn jedes Bild muss nicht nur gedeutet, es muss zunächst ‚gebildet‘ werden: Erst durch dieses weitere Einwohnen wird die assoziative Reichweite der Titelmetapher sichtbar, es eröffnen sich neue Vorstellungszusammenhänge, die dazu führen, die von der vordergründigen Übertragung verdrängten Aspekte aufzuspüren. Genutzt wird hier das Wissen des Metaphernanwenders im Herkunftsbereich seiner Metaphern. Dieses Kenntnisgefälle zwischen Herkunftsbereich und Zielbereich ermöglicht ein besseres Verstehen des zu erforschenden Gegenstandes. Diese zweite Ebene zeigt hier auch den Interaktionsprozess zwischen Bildspender und Bildempfänger. Die gehörte Musik lässt sich dabei nicht wegblenden, es geht hier nicht ausschließlich um die Beschreibung einer Kurve, sondern der gehörte Symphoniesatz schwingt in den Schüleräußerungen immer mit und beeinflusst die sprachliche Ausgestaltung der Metapher. Dieser zweite Schritt des metaphorischen Prozesses ist die eigentlich entscheidende Phase des Unterrichts, weil hier Sprachmaterial bereitgestellt wird, Formulierungen für die eigene musikalische Analyse. Die Schüler sind im Unterricht keinesfalls sprachlos, sondern haben ein sehr detailliertes Sprachvermögen in den Konzeptbereichen ihrer Metaphern. Sie arbeiten nicht mit Material, das von außen herangezogen wird, sondern das sich direkt an der Metapher und damit unmittelbar ästhetisch Wahrgenommenen entzündet:

*„Das Stück beginnt mit einer Vollbremsung.“
„Man ist orientierungslos.“
„Die Musik ist ein Bremsklang.“*

Dieses Einwohnen in die Metapher eröffnet neue Perspektiven, lenkt den Blick auf Phänomene, die über den substitutionstheoretischen Übersetzungsversuch weit hinausreichen. Hier sieht man auch deutlich, wie im Zuge der Ausgestaltung einer Metapher neue Gegenstandsaspekte eröffnet werden und die Strategien der musikalischen Konzeptionalisierung blitzschnell wechseln:

32 Konersmann 2006, S. 14.

Aus einer Raum- oder Landschaftsmetapher („Ein kurviger Weg“) wechselt die Konzeptionalisierung in den Bewegungs- und energetischen Bereich („Die Musik ist ein Bremsklang“). Außerdem werden Momente zur Sprache gebracht, die Allgemeines ansprechen und über die eigentliche Beschreibung des musikalischen (Form)Sinns hinausgehen und den musikalischen Gehalt implizieren, zumindest aber zu seiner Deutung führen können. Ein Beginn, der die Orientierung versagt, wird für die Schüler zum Appell, dieser Konstellation nachzugehen, dabei eigene Erfahrungen einzubinden:

*„Keine Autofahrt beginnt mit einer Vollbremsung.“
„Man beginnt nicht gleich mit einer Kurve, einem Umweg.“*

Der symphonische Start mit einer Kurve lässt Fragen entstehen, die die historische Dimension in assoziative Reichweite geraten lassen und den Rezenten einer Aufführung aus dem Jahr 1805 zu Wort kommen lassen, der ebenfalls aus der Kurve geraten ist:

„... werde hier nur noch angemerkt, dass die Wahl beider Symphonien, welche die Concerte eröffneten, eben nicht ganz glücklich war. Die erste von Beethoven hub in dem Augenblick, als ein gespanntes Publikum die erste kräftige Zusammenstimmung eines grossen zahlreichen Orchesters erwartete, mit dem Septimenakkorde über der Dominante des Haupttons auf einem kurzen Auftakte an. Dergleichen Freiheiten und Eigenheiten wird niemand an einem genialischen Künstler wie Beethoven tadeln, aber ein solcher Anfang passt nicht zur Eröffnung eines grossen Concerts in einem weiten Operntheater.“³³

Beethovens Sinfoniesatz beginnt weder mit einem Septimenakkord auf der Dominante, noch ist er auftaktig, wie es der Berichterstatter hier annimmt. Doch wer seine Tonwanderung in einer Kurve ohne festen Bezugspunkt beginnt, weiß nicht, aus welcher Richtung er kommt und wohin die Reise führt.

Das spontan assoziierte Bild, im zweiten Schritt verfeinert und verdichtet, soll nun im dritten Schritt an den ursprünglichen Gegenstand angebunden werden. Die letzten Schüleräußerungen zeigen, wie die metaphorischen Mo-

33 Berlinsche Musikalische Zeitung, Jg. 1, 1805, S. 7. Zit. nach Kunze 1987, S. 22. Dieses Zitat war kein Bestandteil des protokollierten Unterrichts. Metaphern eröffnen offene Prozesse. Wer mit einer Metapher beginnt, weiß nicht, wohin die Reise führt.

mente nun (Fach-)Begriffliches exemplifizieren können: „Die Übertragung, die das konkrete Bild der Metapher darstellt, wird – nach seiner ausweitenden Verdichtung – auf die übergreifende theoretische Anschauung ‚gegenübertragen‘, retransferiert. Die Metapher erfüllt dabei faktisch die Funktion eines ‚Gedankenexperiments‘.“³⁴ Jain spricht in diesem Zusammenhang von „Rückübertragung“, diese Möglichkeit wird dabei durch die Betonung des überschreitenden Moments zugleich eingeschränkt. Da es sich hier gerade nicht um einen substitutionstheoretischen Übersetzungsprozess handelt, sollte hier vielleicht besser von Anbindung gesprochen werden. Diese Anbindung ersetzt nicht die Metapher, die den Gegenstand quasi umstellt, hier bilden sich vielmehr Kontaktfäden zwischen dem Gegenstand und seiner metaphorischen Umschreibung. Das metaphorische „sehen als“ wird also nicht durch ein rationales „sehen dass“ ersetzt, vielmehr bilden beide Elemente ein fein abgestimmtes Zusammenspiel. Dies ist der Ort, an dem die Metapher als subjektive Sinnzuschreibung, also das, was als unvermuteter Fund aus dem Dunkel ans Licht gedrungen ist, ihre Rationalität entfalten muss, der Überprüfung am Gegenstand standhalten muss. In der Begrifflichkeit der Schüler begegnen wir nun dieser Anbindung an die Anschlussmetapher „Die Musik ist ein Bremsklang“:

*„Die Lautstärke bricht zusammen.“
 „Die Streicher beginnen verdeckt, abgerissen mit Pizzikato.“
 „Jeder Einsatz endet abrupt, erschreckt.“
 „Das Stück hat mehrere Anfänge.“*

Diese Äußerungen zeigen, wie Schüler mit ihren sprachlichen Mitteln ihrem Verstehen Ausdruck verleihen können. Ob jetzt Töne ‚getupft‘ oder ‚staccato‘ gespielt werden, ist zunächst nicht mit einem Verstehenszuwachs verbunden. Dennoch finden sich hier neue initiale Metaphern, die weiteres Erkenntnispotential in sich bergen, eine Arbeit an der Fachbegrifflichkeit herausfordern, die hier nur kurz angedeutet werden soll: Was bedeutet es, wenn die Streicher verdeckt spielen? Schlüsse erschrecken sind?

*„fallende Dynamik“
 „fp: keine Klänge, die natürlich verklingen, sondern ausgebrems,
 abgedrosselt, erwürgt werden“
 „pizz: Klangstöße, die sofort erlöschen“*

34 Jain 2002, S. 54.

Nun handeln die expliziten Eigenschaften quasi selbst, ohne den Zwischenschritt einer metaphorisch benannten Szene oder handelnden Person. Die Metapher ist mehr als ein Transportmittel. Durch die vorgenommene metaphorische Konzeptionalisierung durchwebt sie das begriffliche Sprechen und vermag es auch, eine Strukturanalyse durch Momente der ursprünglichen ästhetischen Erfahrung anzureichern, die die Musik in ihrer gesamten analogen Gestalt be- und ergreift: „Sie konserviert den Reichtum ihrer Herkunft, den die Abstraktion verleugnen muß.“³⁵ Dabei haben sich die „Bremsklänge“ schon auf den Weg zu einem fixierten Begriff gemacht. Dies entspricht jenem Abnutzungsprozess, dem jede innovative Metapher im unterrichtlichen Diskurs ausgesetzt ist. Sie ergibt sich aus dem Erfahrungskontext des Hörers, eröffnet als Reflex des Erkennens eine erste begriffliche Annäherung, bringt musikalische Formung zum Vorschein.

Auch jüngere Schüler sind in der Lage sich sehr differenziert in ihren Sprachbildern zu äußern. Die Qualität ihrer Bilder steht denen der älteren Schüler in nichts nach. Problematisch bleibt hier die dritte Stufe des metaphorischen Prozesses, die Konsolidierung. Dies gilt übrigens auch, wenn sie aufgrund ihrer individuellen Vorerfahrung – etwa durch Instrumentalunterricht – bereits ein sehr differenziertes begriffliches Inventar zur Verfügung hätten. Sie bleiben in ihrer Bildlichkeit gefangen und stellen fest: „Weil die Musik so ist.“ Aufgabe des Musikunterrichts muss hier also sein, die Kontaktfäden zwischen den Metaphern und dem musikalischen Gegenstand herzustellen oder zu intensivieren.

Der initialen Metapher „ein kurviger Weg“ soll nun eine Analyse von Wilhelm Seidel, der sich einer ähnlichen Konzeptionalisierung bedient, gegenübergestellt werden und sich damit einer etablierten Fachsprache stellen. Dies soll nicht suggerieren, dass nun abschließend ein zu erreichendes Endprodukt aufgelegt wird. Die Metapher ist kein didaktisches Mäntelchen, das nun abgelegt werden kann. Seidel bedient sich einer ähnlichen Konzeptionalisierung, seine Worte konservieren das Metaphorische und tragen die initiale Metapher „ein kurviger Weg“ als assoziative Zwischenschicht förmlich in sich:

„Im Trugschluß, in der Tonikaparallele, wird das Ende des Weges bereits von Ferne sichtbar und der Weg als Umweg fühlbar. Die fallende Dynamik der beiden ersten Akkordpaare vermehrt die Unsicherheit über die Richtung, die die musikalische Bewegung

35 Blumenberg 2001, S. 196.

einschlagen wird.“ Erst im weiteren Verlauf gewinnt die Musik Kontur: „Dort setzen die Violinen, nun streichend, ein und leiten in Achteln zur zweiten Phrase über. Beethoven grundiert sie durch eine stetige Achtelbewegung. Die Musik faßt hier endlich Tritt.“³⁶

Die Metapher bleibt – einmal in die Welt gesetzt – das Ergebnis einer Überschreitung und als Ausdruck unserer unvollkommenen begrifflichen Sprache bestehen. Sie umstellt den Gegenstand: Das metaphorische „sehen als“ kann also nicht durch ein „sehen dass“ ersetzt werden, weil damit ihr Mehrwert entfiere. Gerade dieser Mehrwert spiegelt aber jene subjektive Betroffenheit wider, benennt innere Vorgänge. Verzichtet man auf diesen Mehrwert der Bilder, reißt der Kontakt zwischen Subjekt und Objekt. Das, was der Hörer in den musikalischen Gegenstand als seine Deutung hineingelesen hat und für ihn den eigentlichen Mitteilungswert ausmacht, verschwindet dann hinter den stabilen Gewissheiten einer fachbegrifflichen Ordnung.

Rückblick auf Methodisches

„Die Metapher ist ein einmaliges Sprachelement“³⁷, stellt bereits Werner Ingendahl fest. Einmaliges und System vertragen sich nicht. Einmalige Sprachelemente entziehen sich sowohl einer Quantifizierung, als auch einer Kategorisierung in (zuvor) festgelegte Vorstellungsschemata. Eine individuelle Metapher kann ihr Wissen nicht preisgeben, wenn sie mit einem Lendenschurz an Kontext versehen von der ursprünglichen Unterrichtssituation losgelöst wird und in anonymisierten Wortsammlungen ihre Schärfe einbüßt. Diesen Versuch einer qualitativ systematischen Metaphernanalyse vollzieht bereits Günter Kleinen,³⁸ er gliedert seinen Textkorpus in die vier Kategorien ‚musikalisches Denken‘, ‚sprachliche Metaphern‘, ‚allgemeine Vorstellungsschemata‘ und ‚körperliche Empfindungen‘ und belegt in seiner Auswertung, dass „jegliches Denken in Musik [...] einschließlich der wichtigsten, die Musik betreffenden Metaphern [...] vermittelt über allgemeine Vorstellungsschemata auf körperli-

36 Seidel 1979, S. 8.

37 Ingendahl 1997, S. 305.

38 In Kleinen 1994 anhand ausgedehnter Hörstunden mit 30 Studierenden, in einer weiteren Untersuchung werden bewusst freie sprachliche Äußerungen von 118 Probanden ausgewertet, die keiner Einschränkung durch methodische Vorgaben unterliegen. [Vgl. hierzu Kleinen 1999, S. 55.]

che Empfindungen zurückgeht.“³⁹ Gesucht wurde hier nun nach einem methodischen Zugang, der über diese grundsätzlichen Ergebnisse hinausführen und detailliertere Einblicke in Schülersprache und die damit verbundenen Verstehensprozesse gewähren kann. Ziel ist es dabei, den subjektiven und einmaligen Wahrnehmungseindrücken der Schüler nachzugehen und ihre Relevanz für eine Struktur- bzw. Ausdrucksanalyse zu entfalten. Eine festsegmentierte Kategorisierung einzelner Schülerbeiträge erlaubt keinen Einblick in diese Verstehensprozesse. Diese sind an Kommunikation gebunden. Auch unser Beispiel hat gezeigt, wie im Verlaufe des Verstehensprozesses die Konzeptualisierung oft blitzartig wechselt, wenn Metaphern ihr schöpferisches Potential noch verborgen halten und sich erst im Laufe der Ausdifferenzierung einer Metapher neue Gegenstandsaspekte eröffnen. Diesem Problem widmet sich auch Günter Kleinen mit Unbehagen, wenn er der Ambivalenz seiner qualitativen Untersuchung nachgeht: „So erweist sich, daß die Inhaltsanalyse immer im Fluß bleibt und niemals völlig abgeschlossen ist. [...] Trotz der statistischen Berechnungen bleibt eine Menge an Subjektivität zurück.“⁴⁰ Aber auch die Beschränkung auf eine Auswertung exemplarischer Fallbeispiele ist nicht frei von Vorannahmen, willkürlicher Auswahl oder Akzentuierung einzelner Schüleraussagen in einem komplexen Unterrichtsgeschehen: „Die Metapher ist ein Versprechen, das erst vom Interpreten eingelöst wird.“⁴¹ Als Interpreten fungieren hier nicht nur Schülerinnen und Schüler, hier stellt sich auch der auswertende Beobachter mit seinen eigenen subjektiven Sinnkonstruktionen als aktiver Gefährte in das Geschehen: „Wir sind die Komplizen des metaphorischen Geschehens, unsere Rezeptionskompetenz und Kombinationsgabe vollendet den Effekt.“⁴² Die Gesetzmäßigkeiten, nach denen Metaphern verstanden werden, sind leitend für die Methode der Analyse einzelner Fallbeispiele: „[Metaphern] versorgen [...] den erkennenden nicht nur mit einer Landkarte – sagen wir: mit einer Landkarte des Wissens – , sondern geben auch die Richtlinien vor, wie diese Landkarte zu erstellen sei.“⁴³ Eingebunden wird die individuell gebildete Metapher in eine Metaphorologie im Sinne Hans Blumenbergs⁴⁴. Eine am Phänomen der Metapher orientierte Methode, die

39 Kleinen 1999, S. 67.

40 Kleinen 1999, S. 67.

41 Fermandois 2003, S. 432.

42 Konersmann 2007, S. 14.

43 Konersmann 2007, S. 16, hier in Anlehnung an Thomas S. Kuhn.

44 Vgl. Blumenberg 2001.

nicht allein auf eine Theorie metaphorischer Sprache abzielt, sondern diese vor dem Hintergrund der Theoriebildung untersucht.

Risiken und Nebenwirkungen: Die Klugheit der Metapher in der Dualität von Gefahr und Sicherung

Bei genauerer Betrachtung der Metaphernkommunikation wird auch klar, dass die Metapher in vielen Fällen nicht die Lösung, sondern das Problem im Vermittlungsprozess darstellt. Daher muss es hier auch darum gehen, Risiken und Nebenwirkungen der Metaphernkommunikation zu verhandeln. Das Gefühl, die richtigen Worte gefunden zu haben, stellt sich nicht immer ein. Oft zeigen Metaphern ihre Leuchtkraft gerade durch die metaphorische Differenz, durch ausgewiesene Bildbrüche. Unbestimmtheitsstellen bleiben offen, Metaphern werden ganz zurückgezogen, es entstehen Anschlussmetaphern, die andere Momente des Gegenstandes treffen. Was ist, wenn der Verfasser von seinen Metaphern fehlgeleitet wird, wenn sie keiner Überprüfung standhalten? Oder entzieht sich ein konstruktivistisches Gegenstandsverständnis dieser Metaphernkontrolle? Kluge Metaphern mit zu viel Respekt zu behandeln kann auch dazu führen, sich den Suggestionen, der Logik des Bildes hinzugeben und die Logik der Sache aus dem Blick zu verlieren. Auch die Bildersprache führt zu optischen Täuschungen, wenn man außer Acht lässt, wie eine Metapher den Blick auf die Wirklichkeit einengt, verzerrt und zu vorschnellen Schlussfolgerungen führt: Die Metapher verbindet mit einer „Suggestion von Sicherungen, die sie nicht gewinnen kann. Als Erklärung erscheint, was doch nur Konfiguration ist. Die Funktion der Metapher wird aus dieser Dualität von Risiko und Sicherung begreiflich. Sie nutzt die Suggestion der Anschaulichkeit und ist dadurch nicht nur Vorstufe oder Basis der Begriffsbildung, sondern verhindert sie auch in Richtung ihrer Suggestionen.“⁴⁵ Denkwänge durch Analogiebildung entstehen immer dann, wenn man einer Analogie zu sehr folgt, die Begrenztheit einer Berührung zwischen einem Gegenstand der ihn beschreibenden Metapher nicht zur Kenntnis nimmt. Alternative Metaphern können hier herangezogen werden. Oft wird allerdings der Gegenstand gerade durch diese Unschärfe eines Bildes, durch die metaphorische Differenz sichtbar.

Wie das Beispiel gezeigt hat, bedürfen Metaphern der Erläuterung und Verdichtung durch ihre Schöpfer *und* Interpreten. Lichtenbergs Aphorismus „die Metapher ist klüger als ihr Verfasser“ fasst dabei nur die halbe Wahrheit:

45 Blumenberg 1971, S. 212.

Die Metapher ist nicht nur klüger als ihr Verfasser, sie ist auch klüger als der Leser: „Die Klugheit der Metapher [...] steht und fällt mit der Klugheit ihrer Schöpfer *und* Interpreten, denn diese sind es, die die Metapher verwenden, die sie verstehen oder mißverstehen, die sich von ihrer Suggestionskraft fangen lassen oder durch kritische Reflexion ihre Evidenz und Geltung bestimmen.“⁴⁶ Ehrenforth spricht von der nötigen Sprachphantasie, inzwischen haben wir Metaphernkommunikation aus dem Bereich der Sprache gelöst und als kognitiven Prozess verstanden. Es geht also beim Verstehen einer Metapher darum, sich von den eigenen etablierten Denkmustern zu lösen und sich blitzschnell in die Konzeptionalisierungen der Schüler, die sich hinter dem sprachlichen Gewand verbergen, hineinzusetzen und sie hinsichtlich ihrer Tauglichkeit zu prüfen.

Die Grenzen zwischen Metapher und Begriff sind fließend, alles Sprechen über klingende Phänomene befindet sich in einer Schwebelage von Metapher: Töne unterschiedlicher Frequenzen ordnen wir im virtuellen Raum an, sprechen von ‚hoch‘ und ‚tief‘, während Schüler auch von ‚hellen‘ und ‚dumpfen‘ Klängen sprechen. Auf dem Klavier liegen die hohen Töne ‚rechts‘, auf dem Cello ‚unten‘. Dem Laien mag metaphorisch erscheinen, was dem Fachmann als fixierter Terminus geläufig ist. Gerade wenn ein Verstehensprozess angestrebt wird, bei dem sich – spätestens in der Phase der Anbindung der verdichteten Metapher an den Gegenstand – metaphorische Beschreibung und die Sprache der Begriffe und Fachtermini begegnen, die selbst wiederum „tote“ Metaphern sind: Einerseits ist die Beachtung von metaphorischen Ursprüngen in einem pädagogischen Vermittlungsprozess gerade dort von Bedeutung, wo Begriffe nicht (mehr) Metaphern sind. Andererseits besteht die Gefahr, dass sich Metaphern bereits während des Unterrichtsprozesses begrifflich etablieren und sie den Status einer Metapher verlieren. Damit eine Metapher als solche gelten kann, ist es aber gerade konstitutiv, dass sie nicht wahr ist. Als etablierter Begriff verwässert sie nicht nur die tradierte Fachsprache, sondern verliert auch ihr Gefolge von Assoziationen und damit ihre Sprengkraft. Die innovative Metapher „Bremsklang“ erscheint dann den Schülern im Verlauf eines Unterrichtsgesprächs schnell als vermeintlicher Bestandteil einer Fachsprache. Als Begriff, der sich gerade durch seine Anschaulichkeit fester in den Erfahrungsschatz der Schüler eingräbt, als angelernte Termini einer (beim Lehrer) etablierten Begriffssprache.

46 Debatin, 1995, S. 342.

Auch die Metapher ist nur ein begriffliches Korrelat einer inneren Vorstellung, die präsent ist, bevor sie in Sprache gefasst ist: Der Verfasser ist dabei oft klüger als seine Metapher, wenn er abwägt, welche Aspekte einer inneren Vorstellung er im sozialen Gefüge einer Zwangsgemeinschaft „Unterricht“ öffentlich machen möchte. Sein Mut zur Vermutung ist oft begrenzt: Jede Äußerung ist hier rechenschaftspflichtig gegenüber anderen Ohren, die mithören, gebunden an Konventionen und Erwartungshaltungen ihrer Adressaten, Mitschüler wie Lehrer.

Die subjektive Befindlichkeit und die unmittelbare Hörsituation finden Einlass in eine metaphorische Beschreibung: Die Überlegungen hierzu bleiben im Stadium vager Anmutungen, doch kann die Metapher „ein kurviger Weg“ auch als ein Reflex auf die unterrichtliche Analysesituation gelesen werden. Der Schüler ist bereits gestimmt durch jene verschlungenen Pfade langwieriger und anstrengender Analyseprozeduren, die dazu beitragen, ästhetisch Verstandenes in begrifflich nicht zu Verstehendes zu transformieren und sich auch in der spontanen Assoziation „ein kurviger Weg“ niederschlagen können.

Nicht immer ist die Metapher klüger als der Leser. Häufig ist der Leser klüger als der Verfasser und seine Metaphern: Wo beginnt der Interpret einer Schüleräußerung etwas in eine Metapher hineinzulesen, was vom Urheber nicht gemeint ist? Lichtenberg führt seinen Aphorismus weiter: „Wer Augen hat, der sieht in allem.“⁴⁷ Ist diese Auslegung verwerflich? Oder beginnt gerade hier der Lernprozess, wenn ich einem Schüler die Klugheit seiner eigenen Metaphern „vor Augen stelle“ (Um hier ein letztes Mal eine der Lieblingsmetaphern des Aristoteles’ zu benutzen)? Die Metapher kann fassbar machen, wie die Schüler in ihren Bildern sich selbst voraus sind. Dazu bedarf es eines Unterrichts, der den Schüler nicht dort abholt wo er steht, sondern ihn behutsam dorthin begleitet, wo ihn die Klugheit seiner Metaphern bereits hingeführt hat.

Literatur

- ARISTOTELES. 2006 [335-334 v. Chr.]. *Poetik. Griechisch/Deutsch*. Bibliogr. erg. Ausg. 1994, [Nachdr.]. Stuttgart: Reclam.
- BALDAUF, C. 2000. Sprachliche Evidenz metaphorischer Konzeptionalisierung. Probleme und Perspektiven der kognitivistischen Metapherntheorie

47 Lichtenberg 1968, F 369.

- im Anschluss an George Lakoff und Mark Johnson. In: ZIMMERMANN, Ruben, Hrsg. *Bildersprache verstehen. Zur Hermeneutik der Metapher und anderer bildlicher Sprachformen*. München: Fink. (Übergänge; Bd. 38), S. 117–132.
- BLUMENBERG, H. 1971. Beobachtungen an Metaphern. *Archiv für Begriffsgeschichte*, Bd. 15. Bonn: Bouvier, S. 160–214.
- BLUMENBERG, H. 2001. *Ästhetische und metaphorologische Schriften*. Hrsg. von HAVERKAMP, A. Frankfurt a. M.: Suhrkamp.
- DEBATIN, B. 1995. *Die Rationalität der Metapher. Eine sprachphilosophische und kommunikationstheoretische Untersuchung*. Berlin, New York: de Gruyter.
- DEBATIN, B. 2005. Rationalität und Irrationalität der Metapher. In: FISCHER, Hans R., Hrsg. *Eine Rose ist eine Rose... Zur Rolle und Funktion von Metaphern in Wissenschaft und Therapie*. Weilerswist: Velbrück, S. 30–47.
- DREWER, P. 2003. *Die kognitive Metapher als Werkzeug des Denkens. Zur Rolle der Analogie bei der Gewinnung und Vermittlung wissenschaftlicher Erkenntnisse*. Tübingen: Günter Narr Verlag.
- EHRENFORTH, K. H. 2001. Lebenswelt - das "wirklich Erste". Musikpädagogik zwischen ästhetischer Autonomie-Idee und dialogischem Bildungskonzept. In: EHRENFORTH, Karl H., Hrsg. *Musik - unsere Welt als andere. Phänomenologie und Musikpädagogik im Gespräch*: Königshausen & Neumann, S. 33–58.
- FERMANDOIS, E. 2003. Kontexte erzeugen. Zur Frage nach der Wahrheit von Metaphern. *DZPhil*, 52 (3), S. 427–442.
- GANSEN, P. 2009. Ein neues „Plädoyer für Metaphern“. Forschungsstand und Perspektiven erziehungswissenschaftlicher Metaphernforschung. *Vierteljahresschrift für wissenschaftliche Pädagogik* (3), S. 324–345.
- HERZOG, W. 1983. Plädoyer für Metaphern. Versuch, ein vergessenes pädagogisches Problem in Erinnerung zu rufen. *Vierteljahresschrift für wissenschaftliche Pädagogik* (1), S. 299–332.
- INGENDAHL, W. 1971. *Der metaphorische Prozess. Methodologie zur Erforschung der Metaphorik*. Düsseldorf: Schwann. (Sprache der Gegenwart; 14).

- JAIN, A. K. 2002. *Medien der Anschauung. Theorie und Praxis der Metapher*. München: edition fatal.
- KLEINEN, G. 1994. *Die psychologische Wirklichkeit der Musik. Wahrnehmung und Deutung im Alltag*. Kassel: Bosse. (Perspektiven zur Musikpädagogik und Musikwissenschaft; 21).
- KLEINEN, G. 1999. Die Leistung der Sprache für ein Verständnis musikalischer Wahrnehmungsprozesse. In: BEHNE, Klaus-Ernst, Hrsg. *Wahrnehmung und Rezeption*. Göttingen: Hogrefe. (Jahrbuch der Deutschen Gesellschaft für Musikpsychologie; 14), S. 52–68.
- KOHL, K. 2007. *Metapher*. Stuttgart [u.a.]: Metzler.
- KONERSMANN, R., Hrsg. 2007. *Wörterbuch der philosophischen Metaphern*. Darmstadt: Wiss. Buchges.
- KUNZE, S., Hrsg. 1987. *Ludwig van Beethoven. Die Werke im Spiegel seiner Zeit; gesammelte Konzertberichte und Rezensionen bis 1830*. Laaber: Laaber.
- KURZ, G. 2004. *Metapher, Allegorie, Symbol*. 5., durchges. Aufl. Göttingen: Vandenhoeck & Ruprecht.
- LAKOFF, G. & JOHNSON, M. 2004. *Leben in Metaphern. Konstruktion und Gebrauch von Sprachbildern*. 4. Aufl. Heidelberg: Carl-Auer-Systeme-Verl.
- LICHTENBERG, G. Christoph. 1968. *Sudelbücher*. München: Hanser. (Schriften und Briefe / Georg Christoph Lichtenberg. Hrsg. von Wolfgang Promies; 1).
- NIERAAD, J. 1977. *"Bildgesegnet und bildverflucht". Forschungen zur sprachlichen Metaphorik*. Darmstadt: Wiss. Buchges. (Erträge der Forschung; 63).
- NIETZSCHE, F. 1988. *Die Geburt der Tragödie*. In: *Sämtliche Werke. Kritische Studienausgabe in 15 Einzelbänden*. Hrsg. von G. Colli und M. Montinari. München: dtv, Bd. 1, S. 9-156.
- PEYER, A. 2006. "Unsere Sprache kann man ansehen als eine alte Stadt: ...". Metaphern als Auslöser von Sprachreflexion. *Der Deutschunterricht* (6), S. 26–35.
- ROLF, E. 2005. *Metaphertheorien. Typologie - Darstellung – Bibliographie*. Berlin, New York: de Gruyter.

- SCHENKER, H. 1990. *Heinrich Schenker als Essayist und Kritiker. Gesammelte Aufsätze Rezensionen und kleinere Berichte aus den Jahren 1891 - 1901*. Hrsg. von H. Federhofer. Hildesheim, Zürich, New York: Olms. (Studien und Materialien zur Musikwissenschaft).
- SEIDEL, W. 1979. *Ludwig van Beethoven, I. Symphonie C-Dur*. München: Fink. (Meisterwerke der Musik; 17).
- WEINRICH, H. 1976. *Sprache in Texten*. Stuttgart: Klett-Cotta.

Musikalisches Erleben und Lernen zwischen Musikschule und Grundschule

Methodenkritische Reflexionen am Beispiel der Evaluation des Programms „Monheimer Modell – Musikschule für alle“

Musikalische Förderprogramme im Schnittfeld von Grundschule und Musikschule haben in den letzten Jahren den Austausch über die Inhalte, Struktur und Gestaltung Institutionen übergreifende Kooperationen angeregt und auch die begleitende wissenschaftliche Forschung vor neue Aufgaben gestellt. Im Sinne einer Methodenkritik werden in diesem Beitrag spezifische Voraussetzungen und Herausforderungen der Evaluation derartiger Angebote problematisiert und exemplarisch veranschaulicht.¹ Dabei werden notwendige Annahmen und Wertvorstellungen der Evaluationsforschung thematisiert und das methodische Vorgehen im Forschungsprozess kritisch reflektiert, um das Problemfeld der Nutzbarmachung und Übertragbarkeit von Ergebnissen aus beauftragter Evaluationsstudien in den freien wissenschaftlichen Fachdiskurs zu eröffnen. Vor allem zwei Aspekte erscheinen in diesem Zusammenhang von Interesse: das Berufsbild der Lehrkräfte und die Perspektive der Schülerinnen und Schüler in musikalischen Förderprogrammen.

1 Musikpädagogische Interessen und Evaluationsforschung

Die vielfältigen Beziehungen zwischen Menschen und Musiken bilden den Gegenstandsbereich der Musikpädagogik, woraus die Ergründung von Prozessen der Aneignung und Vermittlung von Musik als Aufgabe der Disziplin abgeleitet ist (vgl. Kraemer 2007, 42 u. 45; vgl. auch Abel-Struth 1985, 124ff.). Die hierbei möglichen methodischen Zugänge unterscheiden sich nach ihren wissenschaftstheoretischen Positionen mit jeweils eigenen Erkenntnisinteressen.

1 Diese Beispiele sind der Evaluationsforschung zum Programm „Monheimer Modell – Musikschule für alle“ entnommen, die gemeinsam mit Maria Luise Schulten durchgeführt wurde (Schulten/Lothwesen 2009).

sen und Erkenntnismodi, sie treffen sich aber im Anspruch des Verstehens und Verbesserns pädagogischer Prozesse (vgl. Kraemer 2007, 45). Eine zentrale Voraussetzung dafür ist das Erkennen solcher Prozesse als genuin musikpädagogisch. Die von Hermann Josef Kaiser dementsprechend konsequent betonte „musikpädagogische Perspektive“ vereint ästhetische und pädagogische Dimensionen (vgl. Kaiser 1998, 32) und bestimmt den „Begriff ‚musikpädagogisch‘“ als „analytisches Kriterium“, das „erlaubt Maßnahmen, die Menschen im Hinblick auf andere Menschen unternehmen, als musikpädagogische zu qualifizieren [...]“ (Kaiser 1998, 35; vgl. hierzu auch Vogt 1988). Mit dieser Positionierung sind individuelle musikalische Erfahrungen und Prozesse des Lernens und Lehrens erfasst, wie sie auch im Rahmen der Förderprogramme stattfinden. Wie aber werden dort solche Prozesse ermöglicht und entsprechende Erfahrungsangebote eröffnet?

Die diversen Programme und Projekte als Kooperationen von Musikschulen und Grundschulen unterscheiden sich zunächst grundlegend hinsichtlich in ihren institutionell gebundenen Zielsetzungen (vgl. Imort 2007, Schewik-Descher 2007, Bähr/Jank/Schwab 2005, Niessen 2005, Bähr 2001, 7f; Schep-ping 1987). Der Lehrplan Musik für Grundschulen in Nordrhein-Westfalen² setzt als Ziel des Musikunterrichts, im Verbund anderer Fächern die kindliche Persönlichkeit zu fördern und die „Freude und das Interesse der Schülerinnen und Schüler an Musik zu wecken und zu intensivieren, ihre musikalische Wahrnehmungs-, Erlebnis- und Ausdrucksfähigkeit zu entfalten und ihre musikalischen Kompetenzen zu entwickeln.“ (Lehrplan Musik für die Grundschulen des Landes Nordrhein-Westfalen Entwurf, MSW 28.1.2008, S. 5) Dieses Angebot ist an alle Kinder gerichtet. Die Aufgabe der Musikschule als Institution der „Kunst- und Kulturpflege“ (Kraemer 2007, 140) ist selektiv im Sinne einer Begabtenfindung und -förderung, die Ziele u.a. eine Vorbereitung auf ein Musikstudium wie auch ein breites Angebot für alle Altersklassen (vgl. hierzu auch Bastian 1997). Im Falle von Kooperationen können somit „Abgrenzungs- und Koordinationsprobleme“ (Kraemer 2007, 140) zwischen den Institutionen auftreten. Mögliche Problemfelder beziehen sich auf zentrale Fragen der unterschiedlichen Qualifikationen der Lehrkräfte, auf unterschiedliche Anforderungen und spezifische Arbeitsbelastungen in den jeweiligen Kontexten

2 Herangezogen als Vergleichsbasis zum evaluierten Programm, das ebenfalls in NRW angeboten wird.

(vgl. Imort 2007; Kramer, 2007, 140f).³ Die strukturelle Güte der Kooperationsprogramme entscheidet letztlich nicht nur über jene Aspekte sondern mit der inhaltlichen und didaktischen Gestaltung auch über einen mögliche Erfahrungs- und Lernangebote für die Schülerinnen und Schüler sowie die dazu notwendigerweise zu einzufordernden Erwartungen seitens der Lehrenden an die Lernenden.

Evaluationsforschung als „wissenschaftsbasierte Dienstleistung“ (Beywl 2006, 95) ist beauftragt und zweckgebunden; ist also nicht wie freie wissenschaftliche Forschung einem reinen Erkenntnisinteresse verpflichtet, sondern auf das Hervorbringen eines der Praxis förderlichen Nutzens ausgerichtet und damit fähig, zur Optimierung von Strukturen, Prozessen, Bedingungen und Effekten bestimmter Maßnahmen beizutragen (vgl. Stockmann 2004, 2ff). Zu den Aufgaben von Evaluation gehören u.a. die Überprüfung und Bewertung der effizienten und effektiven Umsetzung von Maßnahmen, die Begleitung und Dokumentation der Veränderungen durch Lernprozesse sowie die Unterstützung einer Zieldefinition seitens der Auftraggeber durch Einleitung nachhaltiger Entwicklungsprozesse durch die Bereitstellung entsprechender Ergebnisse (vgl. Kardorff 2006, 65f; vgl. auch Preskill 2008, Sanders 2000). Dazu nötige konkrete Fragestellungen sind verhandelbar zwischen den Evaluierenden und Stakeholdern (den Mittelgebern und allen im Programm beteiligten Gruppen, von der Organisation über die Ausführenden bis zur Zielgruppe) und sollten entsprechend die unterschiedlichen Perspektiven und Erwartungen in der Planung und Durchführung der Evaluation beachten. Reinhard Stockmann (2004, 3f) unterscheidet die hierin eröffneten kommunikativen Möglichkeiten anhand von vier zentralen Zielfunktionen von Evaluation: grundlegend ist die Gewinnung von Erkenntnissen über die Realisierung der Maßnahmen und deren Bedingungen, Akzeptanz und Effizienz als Basis zur Entscheidungsfindung (Erkenntnisfunktion); dabei werden zwangsläufig im Zuge der Datenerhebung die Handlungsweisen und Verpflichtungen der involvierten Personen sowie deren Einhaltung offengelegt (Kontrollfunktion); die Kommunikation dieser Ergebnisse ermöglicht einen Austausch unter den beteiligten Personen über Gelingensbedingungen und kann so Nachhaltigkeit fördern (Lernfunktion); Evaluationsergebnisse können durch die Nachweisbarkeit von Wirkungen

3 Vgl. hierzu auch die in der nmz veröffentlichten Erfahrungsberichte und Diskussionen von im JeKi-Programm lehrenden Musikschullehrkräften in den Ausgaben 10, 11, 12/2009.

und Effizienz auch zur Begründung der Maßnahme verwendet werden (Legitimationsfunktion).

Die Position als externe, nicht mit den Stakeholdern verbundene Evaluierende ermöglicht es in diesem Zusammenhang, eine genuin musikpädagogische Perspektive zu vertreten, die sich von den Interessen der Auftraggeber durchaus absetzen und diese in fachlicher Hinsicht erweitern kann. Es gilt hier, die Balance zu finden und einen wissenschaftlichen Anspruch beibehalten ohne den Forschungsauftrag aus den Augen zu verlieren: Das wesentliche Ziel ist eine Beschreibung der Beschaffenheit (Faktizität) und Geltung (der Einordnung und Bewertung) des zu evaluierenden Gegenstandes. Dies kann unterschiedlich motiviert sein, Stockmann verweist in diesem Kontext auf einen überwundenen Positionsstreit zwischen kritisch-rationalen und konstruktivistisch-emanzipatorischen Ansprüchen (vgl. Stockmann 2004, 12), wobei Wolfgang Beywl die aufklärerisch-ethischen Ansprüche von Evaluation und die gesellschaftliche Relevanz der erzielbaren Erkenntnisse betont (Beywl 2006, 65). Als Grundlage dazu sind empirisch nachweisbare Ergebnisse anerkannt und gefordert; das entsprechende methodische Vorgehen von Evaluation verdeutlicht den Anspruch der Wissenschaftlichkeit als eine „systematische Anwendung empirischer Forschungsmethoden zur Bewertung des Konzepts, des Untersuchungsplanes, der Implementierung und der Wirksamkeit sozialer Interventionsprogramme.“ (Bortz/Döring 2002, 102).

2 Zielsetzungen, Methoden und Ergebnisse

Die Umsetzung des Forschungsauftrags und das so begründete Forschungsdesign entscheiden wesentlich über die zu erwartenden Ergebnisse und deren Reichweite. Uwe Flick benennt als „Referenzpunkte“ von Evaluationsforschung die angewandte Methodik, deren theoretische Grundlagen und die Umsetzung und Nützlichkeit der Ergebnisse, an denen Evaluationen hinsichtlich ihrer Güte zu bemessen sind (vgl. Flick 2006, 14f). Für die Evaluierenden bedeutet dies, schon bei der Planung von Evaluationsstudien eine theoriegesicherte Position einzunehmen und so eine genuin musikpädagogische Perspektive zu vertreten. Da ein wesentliches Merkmal musikpädagogischer Forschung in einer engen Verbindung zu anwendungsbezogenen Fragen besteht (vgl. hierzu Nielsen 2009, vgl. auch Oerter/Bruhn 2005), ist das Verhältnis von Unterrichtspraxis und theoretischem Diskurs zu beachten. Hierin kann eine Überleitung von Ergebnissen aus Evaluationsforschung in musikpädagogische Interessen eröffnet werden. Letzteres blickt auf den Unterricht in musika-

lischen Förderprogrammen als soziale Situation mit mannigfaltigen Interaktionsmöglichkeiten und deren Zielorientierung, didaktische Aufbereitung und methodische Vermittlung fachlicher Inhalte. Das Ziel von Evaluationsforschung in diesem Feld ist konsequenterweise eine vergleichende Einschätzung der *musikpädagogischen* Intentionen, Inhalte, Situationen und Nutzenseffekte derartiger Programme.

Entsprechend wurden zur Evaluation des Programms „Monheimer Modell – Musikschule für alle“ theoretische Grundlagen aus fachlichen Erkenntnissen und Diskussionen gewonnen, um die angestrebten Aufgaben angemessen und auf aktuellem Diskussionsstand zu erfüllen. Diese Grundlagen richteten sich wesentlich auf die handlungsorientierte Ausrichtung des zu evaluierenden Programms und die darin ermöglichten Freiräume des Lehrens und Lernens; zudem wurden Erkenntnisse zur Berufsbildforschung von Lehrkräften herangezogen (vgl. Schulten/Lothwesen 2009, 5f). Das Forschungsdesign kombinierte Elemente formativer und summativer Evaluation, ausgerichtet auf die Begleitung der Vermittlungsprozesse im laufenden Programm und dessen Bewertung als Produkt. Eingesetzt wurden qualitative und quantitative Forschungsmethoden, um die prozesshafte und situative Bedingtheit des Unterrichts als soziale Interaktion abbilden zu können. Methodisch gesprochen heißt dies, unterschiedliche Zugänge offen zu halten, um die Vielfalt sozialer Situation angemessen erfassen und interpretieren zu können. Der situative Aspekt des Unterrichts gewinnt eine große Bedeutung, da mögliche spätere Auswirkungen der Ereignisse oft nicht im Moment abzuschätzen sind. Daraus folgt die methodologische Notwendigkeit, derartige Situationen zu dokumentieren; möglich ist dies über den Einsatz von Protokollen, Foto- oder Videodokumentation und Fotoevaluation. Entwicklungen und Veränderungen über die Laufzeit des Programms sind unter prozessualen Aspekten zu thematisieren. Die Erhebung subjektiver Erlebnisse und Erfahrungen im Lehr-/Lernprozess von Lehrenden wie Lernenden ist eine wertvolle Ressource, die Vermittlungswege und Lerneffekte beleuchtet; hierzu bieten sich Verfahren der Befragung (Fragebogen, Interview) an.

Die im Rahmen der Evaluation durchgeführten Untersuchungen folgten dem angenommenen Forschungsauftrag, das erste Schuljahr des Programms hinsichtlich seines Konzepts, seiner Realisierung und seiner Praxisanforderungen zu überprüfen und zu bewerten. Als Themenbereiche wurden curriculare (Inhalte und Vermittlungsformen), organisatorische (Bedingungen der Unterrichtspraxis) und sozialpsychologische (Motivationen und Erwartungen) Aspekte benannt, konkrete Forschungsfragen wurden entsprechend formuliert

und im Verlauf der Untersuchung erweitert. Um charakteristische Eigenheiten des Programms umfassend beschreiben zu können und möglichen perspektivischen Unterschieden gerecht zu werden, wurden alle daran beteiligten Personen befragt.⁴ Zur Kommunikation erster Ergebnisse wurde eine Zwischenbilanz vorgelegt, die bereits Empfehlungen zur Entscheidungsfindung aussprechen konnte.

Erhebungsart	Personen	Zeitraum	Befragte
Unterrichtsbesuche	Lehrkräfte der Musikschule	September 2007 Oktober 2007 Dezember 2008	3
schriftliche Befragung	Lehrkräfte der Musikschule	November 2007 Dezember 2008	13 (76,47%) 13 (76,47%)
	Eltern	Juli 2008	113
	Schüler/innen	Dezember 2008	311
Gruppendiskussion	Lehrkräfte der Musikschule	Juni 2008 Februar 2009	6 (35,29%) 7 (41,18%)
Kurzinterview	Lehrkräfte der Musikschule	September 2007 Oktober 2007 Dezember 2008	3
	Eltern	Mai 2008	3
	Schüler/innen	Oktober 2007, Mai 2008	10
Interview	Grundschulleitung	Oktober 2007	1
		Mai 2008	1

Abb. 1: Übersicht durchgeführter Untersuchungen in der Evaluation des Programms „Monheimer Modell – Musikschule für alle (MoMo)“ (nach Schulten/Lothwesen 2009, 7).

Im Folgenden werden anhand zweier Beispiele das methodische Vorgehen der Evaluation veranschaulicht und Ergebnisse zur Diskussion gestellt: 1) am

4 Dies waren Lehrkräfte der Musikschule, Schülerinnen und Schüler der ersten Klassen, deren Eltern und Grundschulleitungen; Gespräche mit der Musikschulleitung wurden als informeller Austausch geführt und machten Ergebnisse zeitnah transparent. Einstellungen der Grundschullehrkräfte wurden entgegen einer ursprünglichen Absicht nicht erhoben, da die Grundschullehrkräfte selbst nicht direkt in das Programm einbezogen sind und sich zudem im Verlauf der Evaluation ein Schwerpunkt auf Inhalte und Unterrichtsmaterialien herausbildete.

Berufsbild der Lehrkräfte im Programm und 2) an der Perspektive der Schülerinnen und Schüler.

2.1 Anforderungen im Berufsprofil der Musikschullehrkräfte

Zur Befragung der Musikschullehrkräfte wurden verschiedene Erhebungsmethoden kombiniert: Unterrichtsbesuche schufen eine Beobachtungsgrundlage, auf der Besonderheiten des Programms erkennbar wurden; dokumentiert wurde das Unterrichtsgeschehen mittels Protokollen, Fotos und Videoaufnahmen. Die Auswertung dieser Daten ermöglichte eine fokussierte Prüfung hervorstechender Aspekte in schriftlichen Befragungen⁵; hierbei wurden die qualitativen Beobachtungen der Unterrichtsbesuche in quantifizierbare Daten überführt und Items formuliert bzw. aus vorliegenden Studien u.a. zur Lehrerforschung übernommen und angepasst (Pfeiffer 1994, Dick 2006). Um die hierin geäußerten Einschätzungen nochmals zu prüfen, wurde eine Gruppendiskussion angeschlossen.⁶ Dieses Schema wurde im Untersuchungszeitraum zweimal durchlaufen und erhob als Methodentriangulation qualitative und quantitative Daten, die einander ergänzten und ein umfassendes Bild der Erfahrungen und Einschätzungen der Lehrtätigkeit im Programm vermittelten.

-
- 5 Die erste schriftliche Befragung der Musikschullehrkräfte wurde zu Beginn der Evaluation aus forschungspragmatischen Gründen im online-System survey-monkey.com durchgeführt. Bekannte Schwierigkeiten von online-Studien (Auswahlverfahren der Stichprobe, Nonresponse-Fehler, u.a.; vgl. hierzu Couper /Coutts 2006) traten nicht auf, da die Lehrkräfte vorab im Gespräch informiert und per e-mail eingeladen wurden; die Rücklaufquote lag hier wie in der späteren zweiten, Papier-und-Bleistift-Befragung bei etwa 76%.
 - 6 Dies entsprach auch dem Wunsch der Lehrkräfte nach verstärktem Austausch innerhalb des Kollegiums und wurde als Evaluationsmaßnahme positiv angenommen.

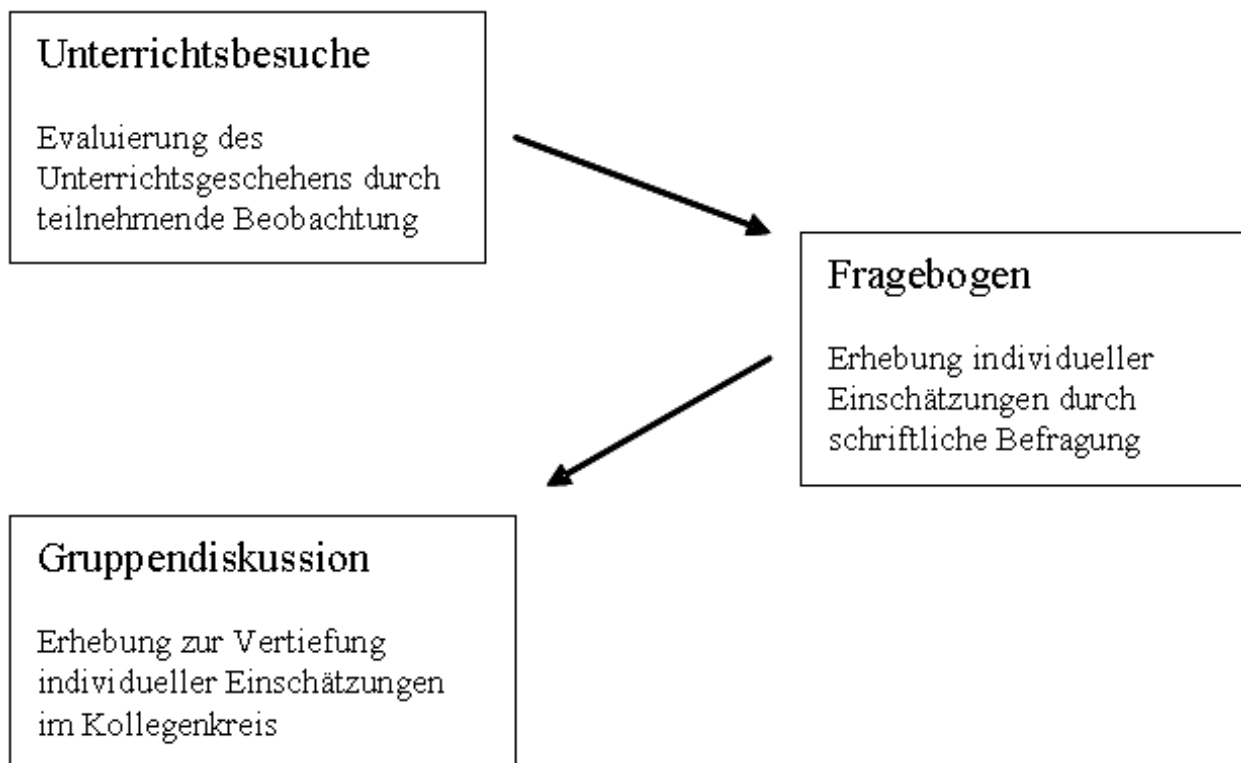


Abb. 2: Modellstruktur der Untersuchungen zum Berufsbild der MoMo-Lehrkräfte (aus Schulten/Lothwesen 2009, 10).

Die Befragung der Lehrkräfte war über nahezu den ganzen Untersuchungszeitraum ausgedehnt, um die Lehrtätigkeit begleiten und etwaige Veränderungen in ihrem prozesshaften Charakter abbilden zu können. Die aus den Ergebnissen profilartig zusammengefassten Eigenschaften der Lehrkräfte decken sich dabei mit Beobachtungen anderer Evaluationsforschungen zu Förderprogrammen (vgl. Beckers/Beckers 2008, 187ff.), sodass diese möglicherweise als Aspekte einer veränderten Berufspraxis anzumerken sind:

- Im Verlauf der Evaluation war ein Prozess der Selbstprofessionalisierung der Lehrkräfte zu beobachten. Dieser sucht eine individuelle Vermittlung zwischen den in der instrumentalpädagogischen Ausbildung erworbenen Qualifikationen und den Anforderungen des Unterrichts im Förderprogramm, besonders der einleitenden Phase mit Inhalten musikalischer Grundausbildung, und wurde vor allem über das kontinuierliche Tun und den Austausch im Kollegium gefördert.
- Die klare Identität der Lehrkräfte als Musikerin bzw. Musiker als zentrale Grundlage des Unterrichtsgeschehens trat ebenfalls zunehmend deutlicher hervor. Diese Selbstbewusstwerdung ermöglichte den Lehrkräften eine klare Positionierung im Hinblick auf den grundschulspezifischen Rahmen ih-

rer Arbeit im Programm und eine flexible Handhabung der vorliegenden Unterrichtsmaterialien.

- Damit einher geht auch eine gesteigerte Kompetenz für die spezifischen Anforderungen des Programms. Die Lehrkräfte unterscheiden deutlich Musikschularbeit und deren Bedingungen von der Unterrichtspraxis im Programm, sie zeigen ein hohes pädagogisches Engagement, wünschen aber zur Qualitätssicherung Fortbildungen zu grundschulspezifischen Fragen (Disziplinprobleme, Stufendidaktik).

Insgesamt zeigten sich Einschätzungen zu Gelingensbedingungen, die den spezifischen Charakter musikalischer Förderprogramme kennzeichnen (vgl. auch Beckers/Beckers 2008, 70ff.). Um diesen Erwartungen und Voraussetzungen näher zu kommen, wurden als Datentriangulation übereinstimmende Themenfelder bei Lehrkräften und Eltern mittels identischer Items abgefragt, betreffend allgemeine Aspekte des Programms sowie die Charakterisierung und Förderung der Schülerinnen und Schüler.⁷

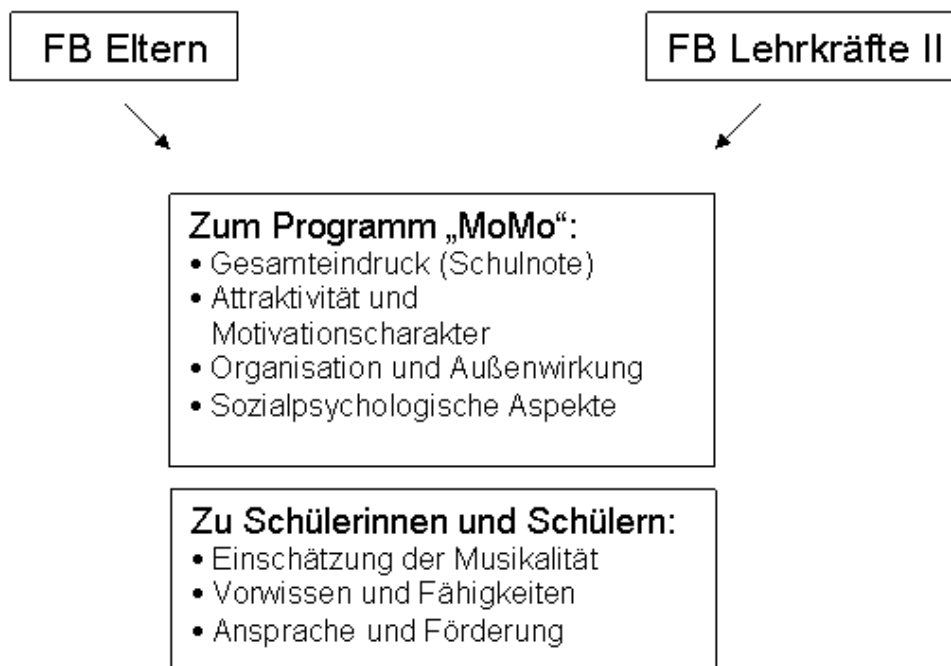


Abb. 3: Übersicht identischer Themenbereiche der schriftlichen Befragungen von Eltern und Lehrkräften.

7 Ein Vergleich des Urteilsverhaltens ist aufgrund der stark unterschiedlichen Stichprobengrößen jedoch nicht gerechtfertigt. So kann hier nur von Tendenzen gesprochen werden, die in weiteren Evaluationen zu prüfen wären.

Das Programm wird von beiden Seiten aus generell positiv bewertet, hinsichtlich möglicher Auswirkungen auf soziale Aspekte (Integration ausländischer Kinder, Sozialverhalten) und vermuteten Transfereffekten urteilen die Lehrkräfte positiver; die Einschätzungen zur Leistungsbereitschaft und -fähigkeit der Kinder fallen hingegen ähnlich aus. Dabei sind jedoch generell die gruppenspezifischen Perspektiven zu beachten: während die Lehrkräfte aus den Erfahrungen der Unterrichtspraxis heraus urteilen, vermitteln die Urteile der Eltern vielmehr erwartungsgebundene Einstellungen (vgl. Schulten/Lothwesen 2009, 25).

2.2 *Musikinstrumente aus Sicht der Schülerinnen und Schüler*

Befragungen von jungen Schulkindern gestalten sich nicht unproblematisch (vgl. Behne 1974; Kleinen 1999). Größere Gruppenbefragungen mittels Fragebogen müssen Einschränkungen antizipieren, die durch die mangelnde Schreib- und Lesekompetenz im ersten Schuljahr bedingt sind. Nach ersten schlaglichtartig eingefangenen Stimmungsbildern im Anschluss an den Unterricht bei denen eine reduzierte Fotoevaluation zum Einsatz kam, wurden alle Kinder der ersten Schuljahre der am Programm teilnehmenden Grundschulen schriftlich befragt.

Der Fragebogen reduzierte sich nach intensiver Auswahl der Themen und Fragen sowie der Klärung der Durchführungsmodalität auf vier geschlossene und eine offene Frage, zzgl. Angaben zu Alter und Geschlecht. Die Befragung wurde von den Musikschullehrkräften während des Unterrichts im Programm durchgeführt. Die Lehrkräfte wurden durch schriftliche Instruktionen über Absicht und Inhalte der Befragung sowie über die Art der Durchführung informiert; der Fragebogen sollte mit den Kindern gemeinsam, Frage für Frage bearbeitet werden.⁸

Die Befragung zielte auf das Involvement der Schülerinnen und Schüler im Programm, d.h. auf den Gesamteindruck der Motivation, des Lernklimas und des Engagements bzw. der Bereitschaft etwas zu lernen. Die Fragen waren einfach gehalten und lehnten sich in der Formulierung an eine kindgemäße

8 Da die Durchführung derart delegiert wurde, können keine Angaben gegeben werden über den Ablauf und mögliche Schwierigkeiten der Befragung. Dies mag forschungsmethodisch als Manko angesehen werden, forschungspragmatisch und angesichts der begrenzten Kapazitäten hinsichtlich Zeit, Personal und Finanzen erschien es im Rahmen der Evaluation vertretbar.

Ausdrucksweise an. Mit einer dreistufigen Smiley-Skala wurden Einschätzungen zu Items erfragt, die als Indikatoren für die gewählten Themen standen. So war bspw. der „Spaß“ am Musikmachen mit Instrumenten als Indikator für die Attraktivität und Motivationsgehalt der musikpraktischen Anteile des Programms gedeutet. Das soziale Gruppenklima wurde mit einer Frage nach Freunden in der Lerngruppe erfasst und als Indikator für emotionale Erregung und Valenz interpretiert. Die Frage nach der Bereitschaft, das im Unterricht erlernte auch „zu Hause“ zu zeigen wurde als Übertrag von Unterrichtsinhalten in den alltäglichen Kontext des Familienlebens verstanden.

Auch sollten Instrumentenpräferenzen erhoben werden, da die Lehrkräfte die Überführung der Kinder in den optionalen Gruppenunterricht am selbst gewählten Instrument ab dem zweiten Schuljahr als implizites Ziel äußerten. Vorliegende Studien arbeiten z.T. mit vorgegebenen Zeichnungen von (überwiegend) Orchesterinstrumenten (vgl. Bähr 2001). Um die Konzentrationsfähigkeit der Schülerinnen und Schüler nicht zu überfordern und möglichst facettenreiche individualisierte Informationen zu erhalten, wurden freie Kinderzeichnungen als projektive Erhebungsmethode eingesetzt (vgl. Reiß 2000, 236f). Dieses Vorgehen wurde in musikpädagogischer Forschung bisher vor allem genutzt, um sozialisatorische Aspekte (Kleinen 1986 u. 1999, Weber/Bullerjahn/Erwe 1999, vgl. auch Hafen 2000) und individuelles Musikerleben (Gotthardt 2006) zu ermitteln. Im vorliegenden Kontext wurde darüber weniger auf die Valenz von Musik als Gegenstand des persönlichen bzw. familiären musikalischen Alltags gezielt als vielmehr auf die individuelle Bedeutung des Affektgehalts von Musikinstrumenten. Um auch emotional bedeutsame Urteile zu ermöglichen, war die entsprechende Frage als Aussage über eine hypothetische Situation in der eigenen Zukunft mit klarer Ich-Bezogenheit übersetzt („Wenn ich groß bin möchte ich dieses Musikinstrument spielen.“) und konnte zeichnerisch oder schriftlich beantwortet werden, mit oder ohne Unterstützung der Lehrkraft. Insgesamt wurden 19 Musikinstrumente genannt, schriftlich oder zeichnerisch; die Zeichnungen wurden hinsichtlich grundlegender Fragen ausgewertet:

1. Wer bzw. was ist dargestellt (Orientierung an Gegenstand, Selbst oder Gruppe)?
2. Wie ist es dargestellt (zeichnerische Qualität, musikbezogenes Wissen und Lerneffekte)?
3. Wie sind Instrumentenwünsche verteilt hinsichtlich des Geschlechts (Genderaspekte)?

ad 1) Orientierungen der Kinderzeichnungen

Im vorliegenden Fall einer klar formulierten Frage waren die Antwortmöglichkeiten zwar sehr offen gehalten, jedoch finden sich als Antworten sehr überwiegend statische Einzeldarstellungen von Musikinstrumenten ohne Spieler. Dargestellt sind auch Prozesse der Klangerzeugung (z.B. Geigenbogen auf den Saiten) oder seltener konkrete Spielvorgänge (Personen am Instrument) abgebildet. Eine derartige Fokussierung auf ein bestimmtes, in der Kommunikationssituation des Malens vertretenes Element der musikalischen Erfahrungswelt (Instrument oder Spielen oder mit anderen Musizieren, ...) konnte Günther Kleinen in einer Auswertung von 5000 Kinderbildern aufdecken (vgl. Kleinen 1999, 75f). Sind mehrere Instrumente abgebildet, so sind sie ohne Bezug zu einander lediglich aufgelistet (vgl. Weber/Bullerjahn/Erwe 1999, 121). In wenigen Fällen sind Zielvorstellungen des eigenen Musikmachens interpretierbar, bspw. wenn Beschriftungen die Zeichnerinnen und Zeichner selbst als Ausübende ausweisen⁹, wenn noch seltener Freunde in die Darstellung musizierender Kinder mit einbezogen werden, wenn eine ich-bezogene Instrumentenpräferenz schriftlich vermittelt wird („Ich möchte gerne trompete schbielen“, FB 162, Junge, 6 Jahre) oder eigene Fähigkeiten nachdrücklich mitgeteilt werden („Ic kan das Slakzock“, FB 167, Junge, 6 Jahre). Der einer stilisierten Blockflöte beigefügte Kommentar („Hajar spielt Flöte“) zeigt eine starke Identifikation des Kindes mit dem Instrument und bildet zudem einen Entscheidungsprozess ab (s.u.).

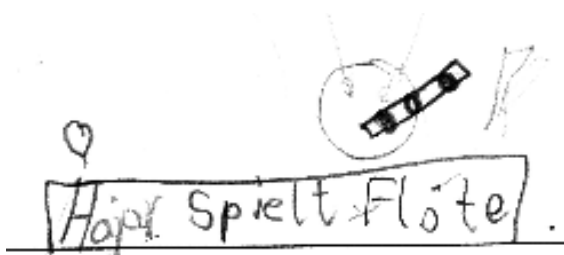


Abb. 4: Kinderzeichnung „Hajar spielt Flöte“ (FB 182, Mädchen, 7 Jahre)

- 9 Claudia Bullerjahn formuliert anhand der Analyse von Kinderzeichnungen eine Abhängigkeit der porträtartigen Selbstdarstellung und persönlichen Beziehung zu den Musikinstrumenten (vgl. Weber/Bullerjahn/Erwe 1999, 121). Diese These könnte auch hier angebracht werden, scheint jedoch nicht hinreichend überprüfbar, da die zum Befragungszeitpunkt im Instrumentenkarussell gesammelten Erfahrungen eigenen Instrumentalspiels nicht erhoben werden konnten. Für nachfolgende Evaluationen wäre hieraus eine hypothesentestende Untersuchung zum Affektgehalt von Musikinstrumenten und Lerneffekten der Instrumentenpräsentation abzuleiten.

ad 2) Imitationslernen, Detailwissen und Entscheidungsprozesse

Aus musikpädagogischer Perspektive sind auch die Lerneffekte einer durchgeführten Fördermaßnahme zu erheben und an deren Inhalten und Vermittlungsarten zu bemessen. Vorliegende Evaluationen haben derartige Veränderungen und Begründungen mit unterschiedlichem methodischem Vorgehen und anhand unterschiedlicher Datentypen aufdeckt, wie z.B. mit dem Einsatz von Testverfahren als Lernkontrolle (vgl. Noll 1992, 172ff, vgl. auch Bähr 2001, 63ff.) oder Interviews mit Schülerinnen und Schülern zur Wissensabfrage über Musikinstrumenten (vgl. Beckers/Beckers 2008, 131f). Da in den Unterrichtsmaterialien des Programms Kopiervorlagen zum Ausmalen vorhanden sind und das Instrumentenkarussell einen großen zeitlichen Umfang im ersten Schuljahr einnimmt, sollte anhand der schriftlichen bzw. zeichnerischen Äußerungen auch der Wissensstand über Musikinstrumente abzulesen sein.

Tatsächlich sind in den Kinderzeichnungen zunächst eine Erweiterung musikspezifischen Wissens und die Fähigkeit, dies mitzuteilen erkennbar (vgl. Hafen 2000, 520f). Darüber hinaus öffnet sich hier ein Blick auf Lerneffekte im Sinne einer zeichnerisch „korrekten“ Wiedergabe der Instrumente, wie auch im Sinne eines musikspezifischen Wissens über Funktion und Bedienung von Musikinstrumenten; der durch Ausradieren, Übermalen und Beschriften nachvollziehbare Zeichenprozess lässt in einigen Fällen auch Präferenzentscheidungen sichtbar werden (vgl. Kleinen 1999, 76). Bemerkenswert ist die häufig detailgetreue Zeichnung, die instrumentenspezifische Eigenheiten berücksichtigt –wenn auch die Größenverhältnisse oftmals nicht den korrekten Relationen entsprechen.¹⁰ Dies ist als Imitationslernen zu beschreiben, das auf das Instrumentenkarussell des ersten Schuljahres und die verfügbaren Unterrichtsmaterialien (Kopiervorlagen zum Ausmalen) zurückzuführen ist; so wird in den Kinderzeichnungen z.B. klar unterschieden zwischen Block- und Querflöte und deren jeweiligen Mundstücken. Die hier gezeigte Gitarre weist er-

10 Günther Kleinen verweist in diesem Zusammenhang auf die psychologische Bedeutung der Größenverhältnisse aus Sicht von Kindern (vgl. Kleinen 1999, 74; vgl. auch Hafen 2000, 525; Reiß 2000). Unterscheidungen hinsichtlich des Präzisionsgrad der Darstellungen, wie von Roland Hafen vorgenommen (vgl. Hafen 2000, 520f), erscheinen hier nicht aussagekräftig, da keine individuellen Leistungsstände und musikalischen Vorerfahrungen der Kinder erhoben wurden und keine entsprechenden Abhängigkeiten geprüft werden konnten (auch Hafen konnte keine Gründe für Unterschiede des Präzisionsgrades aufdecken, vgl. ebd., 521).

kennbar Saiten, Schallloch und Wirbel auf; mitunter ist sogar der Steg erkennbar:

- Wenn ich groß bin, möchte ich dieses Musikinstrument spielen:

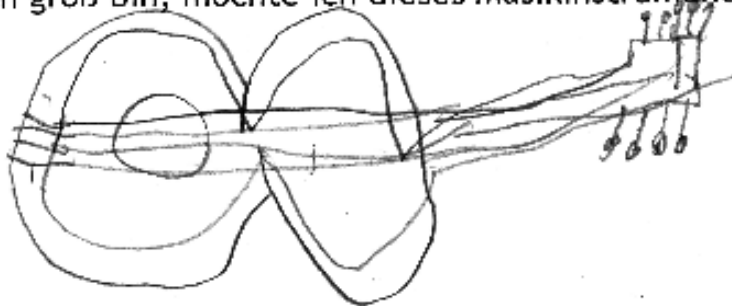


Abb.5: Kinderzeichnung Gitarre (FB 73, Junge, 7 Jahre).

Auch wenn diese vielleicht noch nicht selbsttätig erfahren wurde, ist die jeweils instrumententypische Klangerzeugung ebenfalls aus den Zeichnungen ablesbar: bei Geigen ist ein Bogen angefügt, bei Trommel und Schlagzeug (drumset) zwei Stöcke.

ad 3) Instrumentenpräferenzen und Genderaspekte

Die Ergebnisse hinsichtlich der Präferenz für bestimmte Instrumente spiegeln Befunde vorliegender Studien: während die beliebtesten Musikinstrumente für Jungen Schlagzeug und Gitarre sind, wünschen Mädchen zuvorderst Gitarre, Flöte und Geige (vgl. Scheuer 1984 u. 1988, 202; Stolla/ Gaul 2008, 172).¹¹ Unterschiede sind auch hinsichtlich der nominalen Breite der Nennungen aufzuzeigen: die befragten Jungen nennen mehr verschiedene Instrumente als die Mädchen (darunter eine Eierrassel, ein Schellenkranz und eine Ukulele).¹² Susan Hallam et al (2008) unterscheiden in ihrem Modell der Instrumentenwahl

11 Auch der VdM zählt Gitarre, Flöte und Violine zu den „beliebtesten Instrumenten“ (vgl. URL: <http://www.musikschulen.de/musikschulen/fakten/die-beliebtesten-instrumente/index.html>).

12 Eine statistische Prüfung der Angaben muss dies berücksichtigen. Geordnet nach der dieser Studie zugrunde gelegten Fächerklassifikation des VdM (Zupf-, Streich-, Tasten-, Schlag-, Holz- und Blechinstrumente) sind keine signifikanten Unterschiede feststellbar (Kendalls Tau-b = .333); ausgehend von der ungeordneten Gesamtliste konkreter Instrumentennennungen sind jedoch entsprechend deren breiterer Menge gendertypische Effekte nachweisbar (Kendalls Tau-b = .661, $p < 0.01$).

individuelle, soziale und instrumentenspezifische Faktoren. Entscheidend für den vorliegenden Kontext ist der Zugang zu Musikinstrumenten im Rahmen musikalischer Förderprogramme als sozialer Situation: Einige vorgestellte Instrumente können im Folgejahr nicht im Gruppenunterricht angeboten (z.B. Klavier), andere von den Kindern geäußerte Instrumente evtl. nicht im Instrumentenkarussell präsentiert werden; die Attraktivität von Musikinstrumenten scheint zudem unmittelbar gebunden an die Lehrkräfte. In diesem Moment erhalten gendertypische Aspekte ein stärkeres Gewicht. Studien von Susan O'Neill (O'Neill/Boulton 1996, Harrison/O'Neill 2000) unterstreichen die auch von Scheuer (1988, 203f) festgestellten stereotypen Rollenvorstellungen der Instrumentenwahl. Dabei scheinen weniger konkrete Instrumente als vielmehr ihre Konnotation als typisch weiblich bzw. männlich ausschlaggebend.

Gattung	Instrument	Mädchen			Jungen		
		Anzahl	Summe	Prozent	Anzahl	Summe	Prozent
Zupf			56	30,77		71	41,04
	Gitarre	56			63		
	E-Gitarre	0			7		
	Ukulele	0			1		
Streich			29	15,93		9	5,2
	Geige	29			9		
Tasten			21	11,54		14	8,09
	Klavier	17			10		
	Keyboard	2			4		
	Melodika	2			0		
Schlag			16	8,79		40	23,12
	Schlagzeug	4			23		
	Trommel	11			13		
	Triangel	1			2		
	Eierrassel	0			1		
	Schellenkranz	0			1		
Holz			53	29,12		26	15,03
	Flöte	39			16		
	Querflöte	11			4		
	Saxophon	2			3		
	Blockflöte	0			2		
	Klarinette	0			1		
	Mundharmonika	1			0		
Blech			7	3,85		13	7,51
	Trompete	7			13		
SUMME		182		100	173		99,99

Abb. 6: Aufstellung genannter Instrumente nach Gattungen

Für die Erschließung eines Zugangs zu Musikinstrumenten erscheint nun vor allem die Art der Präsentation relevant, über die derartige Geschlechterstereotype aufgebrochen werden können (vgl. Hallam/Rogers/Creech 2008); den Lehrkräften kommt somit auch in dieser Hinsicht die Funktion eines Rollenmodells zu.

Die Methodenkombination von Beobachtungen im Rahmen der Unterrichtsbesuche, anschließenden Kurzinterviews und einer schriftlichen Befragung hat eine mehrperspektivische Interpretation der einzelnen Ergebnisse ermöglicht. In der Zusammenschau ergibt dies ein Profil der im Programm unterrichteten Schülerinnen und Schüler, das deren Motivation, Zielvorstellungen und Lernfortschritte beschreibt (vgl. Schulten/Lothwesen 2009, 21 u. 24):

- Die Schülerinnen und Schüler zeigten generell eine hohe Motivation, die wesentlich durch die musikpraktischen Anteile des Unterrichts und die Attraktivität der Musikinstrumente genährt ist (vgl. Beckers/Beckers 2008, 181f).¹³
- Das Ziel der Kinder ist das eigene Musikmachen. Die Befragten verwiesen hierbei auch auf ältere bereits im Programm unterrichtete Geschwister und Freunde, die als Referenz bemüht wurden um den Wunsch nach Instrumentalunterricht und bestimmten Instrumenten zu begründen (vgl. Pape 1998).
- Lerneffekte waren zu beobachten in der Verfeinerung musikpraktischer Fertigkeiten und in der Erweiterung musikalischen Wissens sowie dessen verbale und zeichnerische Mitteilung (vgl. Beckers/Beckers 2008, 131ff.).

3 Einsichten für musikpädagogische Evaluationsforschung

Im Rahmen der durchgeführten Evaluationsforschung zum Programm „Monheimer Modell – Musikschule für alle“ konnte die intendierte Abbildung der Perspektiven aller am Programm beteiligter Personengruppen (Lehrkräfte, Schülerinnen und Schüler, Eltern, Grundschulleitungen) geleistet werden. Daraus konnten spezifische Eigenheiten des Programms ergründet werden, die sich konzeptionell und in der Umsetzung von anderen Förderprogrammen ab-

13 90,7% der befragten Schülerinnen und Schüler äußerten, das Musizieren mit Instrumenten gefalle ihnen sehr. Dies wird durch Einschätzungen von Eltern und Lehrkräften bzgl. der Freude der Kinder am Singen bzw. Musizieren unterstrichen: als „sehr gut“ werteten diese 61,95% der Eltern und 58,3% der Lehrkräfte).

heben. Auf der Grundlage der Ergebnisse konnten letztlich gezielte Empfehlungen zur Optimierung formuliert werden, die als Entscheidungshilfen an die Auftraggeber übertragen wurden. Ein ergänzender Abgleich der Ergebnisse der einzelnen Teiluntersuchungen war möglich durch den Einsatz und die Kombination qualitativer und quantitativer Methoden. Gemäß dem Plädoyer Philipp Mayrings für eine Integration dieser beiden Forschungsrichtungen durch einen flexibel erweiterbaren Forschungsplan (vgl. Mayring 2001), ist das Vorgehen entsprechend zu beschreiben. Der offen gehaltene Forschungsauftrag konnte zunächst durch drei zentrale inhaltliche Felder (curriculare, organisatorische, sozialpsychologische Aspekte) ausgefüllt und die Bereitstellung von Empfehlungen zur nachhaltigen Entwicklung des Programms fokussiert werden (vgl. Schulten/Lothwesen 2009, 6). Wie am Beispiel der Untersuchungen zu Anforderungen des Berufsbilds der Lehrkräfte wurden aus der Beobachtung und Beschreibung erster Daten konkrete Fragestellungen für weitere Untersuchungen gewonnen (s.o.). Die dazu eingesetzten Erhebungsverfahren und Instrumente waren jeweils am Gegenstand bemessen und mit Rückgriff auf bereits geprüfte Messinstrumente erstellt; auf den Einsatz rein standardisierter Messinstrumente wurde verzichtet, um spezifischen Eigenheiten des Programms aufzudecken. Eine stete methodische und ergebniskritische Reflexion im Verlauf der Evaluation war notwendig, um dem Prozesscharakter des Unterrichts gerecht zu werden und die Reichweite der jeweiligen (Teil-)Ergebnisse abzuschätzen. So wurden die einzelnen Personengruppen getrennt befragt und ihre Perspektiven zu einigen Aspekten teilweise vergleichbar gemacht. Aufgrund der gravierenden Unterschiede der Stichprobengrößen waren statistische Verfahren zur Hypothesenprüfung jedoch nicht immer sinnvoll anzuwenden. Hinsichtlich einer Prüfung der Verallgemeinerbarkeit der erzielten Ergebnisse wurden letztlich, wie in diesem Beitrag dargestellt, unterschiedliche theoretische Hintergründe bemüht werden. Dabei ergab sich die Schwierigkeit, dass gerade für die möglicherweise zukunftsrelevanten Veränderungen des Berufsbilds von Musikschullehrkräften nur die Studie von Erich und Renate Beckers (2008) direkt vergleichbare Erkenntnisse bietet.¹⁴

Die Nutzbarmachung der Ergebnisse im allgemeinen Forschungsdiskurs der Musikpädagogik schließt sich an dieser Stelle an und fragt zunächst nach

14 Obgleich die Untersuchungen von Arendt (2009) Schewik-Descher (2007), Grosse (2005) und Swanse (2000) Anhaltspunkte für die Anforderungen der Lehrkräfte in Gruppenunterrichtssituationen bieten, sind die Voraussetzungen der Programme anders gelagert.

der Definition der Ergebnisse: Diese müssen dem angenommenen Forschungsauftrag entsprechen und zugleich wissenschaftlichen Ansprüchen wie auch dem aktuellen Erkenntnisstand der Fachdisziplin genügen. Auch das Eigentumsrecht an den erzielten Ergebnissen ist zu beachten; in entsprechendem Übereinkommen mit dem Auftraggeber können Daten und Ergebnisse auch im freien wissenschaftlichen Diskurs präsentiert werden, so bringen diese Reflexionen Aspekte der Unterrichtserfahrungen der Lehrkräfte und der Instrumentenwahl von Kindern in den Fachdiskurs ein. Eine Weiterführung der skizzierten Ergebnisse in nachfolgenden Evaluationsforschungen würde mit einer faktorenanalytischen Auswertung möglicherweise weitere Dimensionen der Einschätzungen und Wirkungen musikalischer Förderprogramme ermitteln und die beobachteten Tendenzen prüfen können. Ein möglicher Schwerpunkt könnte dabei auf den Motivationen der Schülerinnen und Schüler liegen.¹⁵ Auch wäre das Design der hier diskutierten Evaluationsforschung auf eine mögliche Übertragbarkeit auf andere Programme zu prüfen, um Anhaltspunkte für weitere musikpädagogisch orientierte Evaluationsstudien zu bieten.

Literatur

- Abel-Struth, Sigrid (1985). *Grundriß der Musikpädagogik*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Arendt, Gerd (2009). *Instrumentalunterricht für alle? Zur langfristigen Relevanz des Klassenmusizierens und der Notwendigkeit einer Reform des Musikunterrichts*. Augsburg: Wißner.
- Bastian, Hans Günther (Hrsg.)(1997). *Schulmusik und Musikschule in der Verantwortung. Begabungsforschung, Begabtenfindung und Begabtenförderung „von unten“*. Mainz: Schott.
- Bähr, Johannes (2001). *Zur Entwicklung musikalischer Fähigkeiten von Zehn- bis Zwölfjährigen. Evaluation eines Modellversuchs zur Kooperation von Schule und Musikschule*. Göttingen: Cuvillier.
- Bähr, Johannes; Jank, Werner & Schwab, Christoph (2005). Musikunterricht und Ensemblespiel im Rahmen der Kooperation von allgemein bildender Schule und Musikschule. In: *Ensemblespiel und Klassenmusizieren in*

15 Denkbar ist, hierzu das von Veronica O. Sichivitsa (2007) vorgeschlagene und an US-College Studierenden getestete Pfadmodell zu prüfen, das Einflüsse wie u.a. die Unterstützung von Eltern und Lehrkräften und bisherige musikalische Erfahrungen auf die eigene Motivation verbindet.

Schule und Musikschule, hrsg. von Rudolf-Dieter Kraemer und Wolfgang Rüdiger. Augsburg: Wißner, S. 131–152.

- Beckers, Erich & Beckers, Renate (2008). *Faszination Musikinstrument – Musikmachen motiviert. Bericht über die zweijährige Evaluationsforschung zum Bochumer Projekt „Jedem Kind ein Instrument“*. Münster: Lit.
- Behne, Klaus Ernst (1974). Zur Erfassung musikalischer Verhaltensweisen im Vorschul- und Primarbereich. *Musik und Bildung*, 2, 103–108.
- Beywl, Wolfgang (2006). Evaluationsmodelle und qualitative Methoden. In: *Qualitative Evaluationsforschung. Konzepte, Methoden, Umsetzungen*, hrsg. von Uwe Flick, Reinbek bei Hamburg: Rowohlt, S. 92–116.
- Couper, Mick P. & Coutts, Elisabeth (2006). Online-Befragung. Probleme und Chance verschiedener Arten von Online-Erhebungen. In: *Methoden der Sozialforschung*, hrsg. von Andreas Diekmann. Wiesbaden: Verlag für Sozialwissenschaften, S. 217–243.
- Dick, Rolf van (2006). *Stress und Arbeitszufriedenheit bei Lehrerinnen und Lehrern. Zwischen 'Horrorjob' und Erfüllung*. Marburg: tectum.
- Flick, Uwe (2006). Qualitative Evaluationsforschung zwischen Methodik und Pragmatik. Einleitung und Überblick. In: *Qualitative Evaluationsforschung. Konzepte, Methoden, Umsetzungen*, hrsg. von Uwe Flick, Reinbek bei Hamburg: Rowohlt, S. 9–29.
- Gotthardt, Christoph (2006). Kinder malen Konzerte. Musikverarbeitung zwischen Erlebnisreflexion und Selbstvergewisserung. In: *Lehr- und Lernforschung in der Musikpädagogik* (=Musikpädagogische Forschung, Bd. 27), hrsg. von Niels Knolle. Essen: Die Blaue Eule, S. 85–103.
- Grosse, Thomas (2005). *Instrumentaler Gruppenunterricht an Musikschulen. Eine Untersuchung am Beispiel des Landes Niedersachsen*. Augsburg: Wißner.
- Hafen, Roland (2000). „Ich und die Musik“. Kinder malen Bilder zur Musik. In: Bastian, Hans Günther. *Musik(erziehung) und ihre Wirkung. Eine Langzeitstudie an Berliner Grundschulen*. Mainz: Schott, S. 511–533.
- Hallam, Susan; Rogers, Lynn & Creech, Andrea (2008). Gender Differences in Musical Instrument Choice. *International Journal of Music Education*, 26:1, 7–19.

- Harrison, Anna C. & O'Neill, Susan A. (2000). Children's Gender-Typed Preferences for Musical Instruments. An Intervention Study. *Psychology of Music*, 28, 81–97.
- Imort, Peter (2007). Kooperationen allgemein bildender Schulen und öffentlicher Musikschulen. Stand und Perspektiven aus musikdidaktischer Sicht. In: *Zukunft für Musikschulen. Herausforderungen und Perspektiven der Zukunftssicherung öffentlicher Musikschulen*, hrsg. von Thomas Knubben und Petra Scheidewind. Bielefeld: transcript, S. 99–114.
- Kaiser, Hermann Josef (1998). Was heißt „aus musikpädagogischer Perspektive“? In: *Entwicklung und Sozialisation aus musikpädagogischer Perspektive* (=Musikpädagogische Forschung, Bd. 19), hrsg. von Mechthild von Schoenebeck. Essen: Die Blaue Eule, S. 27–40.
- Kardorff, Ernst von (2006): Zur gesellschaftlichen Bedeutung und Entwicklung (qualitativer) Evaluationsforschung. In: *Qualitative Evaluationsforschung. Konzepte, Methoden, Umsetzungen*, hrsg. von Uwe Flick, Reinbek bei Hamburg: Rowohlt, S. 63–91.
- Kleinen, Günter (1986). Kinderbilder als Erhebungsverfahren zur Musiksozialisation im Grundschulalter. In: *Unterrichtsforschung* (=Musikpädagogische Forschung, Bd. 7), hrsg. von Hermann Josef Kaiser. Laaber: Laaber, S. 51–69.
- Kleinen, Günter (1999). Qualitative Verfahren bei der Erforschung musikalischer Lebenswelten von Kindern. In: *Kinder – Kultur. Ästhetische Erfahrungen. Ästhetische Bedürfnisse*, hrsg. von Claudia Bullerjahn, Hans-Joachim Erwe und Rudolf Weber. Opladen: Leske + Budrich, S. 67–82.
- Kraemer, Rudolf-Dieter (2007). *Musikpädagogik. Eine Einführung in das Studium* (2. verb. Aufl.). Augsburg: Wißner.
- Lehrplan Musik für die Grundschulen des Landes Nordrhein-Westfalen Entwurf, MSW 28.1.2008.
- Mayring, Philipp (2001). Kombination und Integration qualitativer und quantitativer Analyse. *Forum: Qualitative Sozialforschung*, 2:1, o.S. URL: <http://www.qualitative-research.net/index.php/fqs/article/view/967/2111>
- Nielsen, Frede (2009). What Is the Significance of Research for Music Education in Practice? On Relations Between the Practice of and the Scientific Approach to Music Education. *Philosophy of Music Education Review*, 17:1, 22–40.

- Niessen, Anne (2005). *Kooperationsmodelle Schule-Musikschule*. BFG Musikpädagogik, Bericht der Arbeitsgruppe, Soest 2005, S. 51–54.
- Noll, Günther (1992). *Musikalische Früherziehung. Erprobung eines Modells*, Regensburg: Bosse.
- Oerter, Rolf & Bruhn, Herbert (2005). Musikpsychologie in Erziehung und Unterricht, in: *Spezielle Musikpsychologie* (= Enzyklopädie der Psychologie, Serie VII, Themenbereich D, Musikpsychologie Bd. 2), hrsg. von Rolf Oerter and Thomas Stoffer. Göttingen et al.: Hogrefe, S. 555–624.
- O'Neill, Susan A. & Boulton, Michael J. (1996). Boys' and Girls' Preferences for Musical Instruments. A Function of Gender? *Psychology of Music*, 24, 171–183.
- Pape, Winfried (1998). Familiäre Einflüsse auf das Erlernen von Instrumenten bei Kindern und Jugendlichen. In: *Entwicklung und Sozialisation aus musikpädagogischer Perspektive* (=Musikpädagogische Forschung, Bd. 19), hrsg. von Mechthild von Schoenebeck. Essen: Die Blaue Eule, S. 111–129.
- Pfeiffer, Wolfgang (1994). *Musiklehrer. Biographie, Alltag und berufliche Zufriedenheit von Musiklehrern an bayerischen Gymnasien. Eine theoretische und empirische Analyse*. Essen: Blaue Eule.
- Preskill, Hallie (2008). Evaluation's Second Act. A Spotlight on Learning. *American Journal of Evaluation*, 29:2, 127–138.
- Reiß, Wolfgang (2000). Zur Produktion und Analyse von Kinderzeichnungen. In: *Methoden der Kindheitsforschung. Ein Überblick über Forschungszugänge zur kindlichen Perspektive*, hrsg. von Friederike Heinzel. Weinheim u. München: Juventa, S. 231–244.
- Sanders, James R. (Hrsg.)(2000). *Handbuch der Evaluationsstandards. Die Standards des „Joint Committee on Standards for Educational Evaluation“*. (2. durchges. Aufl.) Opladen: Leske + Budrich.
- Schepping, Wilhelm (1987). Zur schuldidaktischen Problematik einer zweispurigen Musikunterweisung in allgemeinbildender Schule und Jugendmusikschule. In: *Außermusikalische Musikerziehung* (=Musikpädagogische Forschung, Bd. 8), hrsg. von Günter Kleinen. Laaber: Laaber, S. 133–146.
- Scheuer, Walter (1984). Präferenzen für Musikinstrumente bei Jugendlichen. Eine laufende repräsentative Studie in Hannover. In: *Kind und Musik*

- (=Musikpädagogische Forschung, Bd. 5), hrsg. von Günter Kleinen. Laaber: Laaber, S. 300–315.
- Scheuer, Walter (1988). *Zwischen Tradition und Trend. Die Einstellung Jugendlicher zum Instrumentalspiel. Eine empirische Untersuchung*. Mainz: Schott.
- Schewik-Descher, Wilfried (2007). *Instrumentalunterricht in der Kooperation von Musikschule und Schulmusik. Aspekte einer problematischen Partnerschaft*. Essen: Die Blaue Eule.
- Schulten, Maria Luise & Lothwesen, Kai Stefan (2009). *Abschlussbericht der wissenschaftlichen Evaluation zum Programm „Monheimer Modell – Musikschule für alle“*. URL: http://www.monheim.de/kultur/musikschule/evaluation_momo_abschluss_lang.pdf
- Sichivitsa, Veronica O. (2007). The Influences of Parents, Teachers, Peers and other Factors on Student's Motivation in Music. *Research Studies in Music Education*, 29, 55–68.
- Stolla, Jochen & Gaul, Magnus (2008). Von der Blockflöte zur E-Gitarre. Eine empirische Studie zu Instrumentenvorlieben bei Kindern und Jugendlichen. In: *Musizieren innerhalb und außerhalb der Schule*. (=Musikpädagogische Forschung, Bd. 29), hrsg. von Andreas C. Lehmann und Martin Weber. Essen: Die Blaue Eule, S. 163–178.
- Stockmann, Reinhard (2004). *Was ist eine gute Evaluation? Einführung zu Funktionen und Methoden von Evaluationsverfahren* (=CEval-Arbeitspapiere 9). Saarbrücken: Centrum für Evaluation.
- Schwanse, Ulrike (2000). *Instrumentaler Gruppenunterricht an den Musikschulen des Landes Nordrhein-Westfalen. Wissenschaftlich-empirisches Forschungsprojekt*. Essen: Die Blaue Eule.
- Vogt, Jürgen (1998). Zum Problem des musikpädagogischen Standortes. Reflexionen zum Beitrag „Was heißt ‚aus musikpädagogischer Perspektive‘?“ von Hermann J. Kaiser. In: *Entwicklung und Sozialisation aus musikpädagogischer Perspektive* (=Musikpädagogische Forschung, Bd. 19), hrsg. von Mechthild von Schoenebeck. Essen: Die Blaue Eule, S. 41–55.
- Weber, Rudolf; Bullerjahn, Claudia & Erwe, Hans-Joachim (1999). Musikbezogene Bedürfnisse und die Bedeutung von Musik für Kinder der 90er Jahre. Überlegungen zu einem Forschungsprojekt. In: *Kinder – Kultur. Ästhetische Erfahrungen. Ästhetische Bedürfnisse*, hrsg. von Claudia Bul-

lerjahn, Hans-Joachim Erwe und Rudolf Weber. Opladen: Leske + Budrich, S. 107–129.

„Wie Lehrer lieber lernen“

Eine qualitative Studie über die Rolle von Fortbildungen aus der Sicht von Musiklehrerinnen und -lehrern

Fragestellung

Forschungsleitendes Interesse für die hier vorgestellte Studie war die Frage erstens nach den Wünschen, die MusiklehrerInnen in Bezug auf Fortbildungen haben und zweitens danach, was aus ihrer Sicht von den bisher besuchten Fortbildungen in ihrem Unterricht angekommen ist bzw. was sie umgesetzt haben. Den persönlichen Hintergrund meines Interesses bildeten – neben Beobachtungen als Teilnehmer von Fortbildungen – zum einen eigene Erfahrungen als Referent bei Fortbildungsveranstaltungen verschiedener Anbieter, zum anderen Erfahrungen bei Diskussionen im Rahmen von Planungsgesprächen als Mitglied des AfS-Bundesvorstands, in denen es um die Konzeption von Bundes- und Landeskongressen ging, oder als Referent bei Diskussionen mit der Planungsgruppe für Fortbildungen der Bezirksregierung Köln. Dabei gewann ich den Eindruck, dass das Fortbildungsangebot sämtlicher Anbieter mehr oder weniger zufällig zustande kommt und insbesondere durch die vorhandenen ReferentInnen und deren Schwerpunktthemen sowie durch aktuelle pädagogische „Trends“ bestimmt wird (außerdem natürlich bei bestimmten Anbietern wie Software-Herstellern oder Verlagen durch deren legitimes kommerzielles Interesse). Das sich abzeichnende Bild war das eines unsystematischen Angebots, bei dessen Zustandekommen organisatorische, planerische Fragen in der Regel „aus dem Bauch heraus“ beantwortet zu werden scheinen, oft basierend auf mehr oder weniger erfahrungsgestützten bis vorurteilsbeladenen Einschätzungen hinsichtlich der Zielgruppe, aber ohne vorherige Bedarfsabfrage und ohne eine Evaluation, die – falls überhaupt durchgeführt – über eine kurze Abfrage im Anschluss an Veranstaltungen hinaus geht und eine nachhaltige Wirksamkeit im Blick hat.

Methodische Anlage der Arbeit

Da es sich bei der Frage nach den Wünschen an Fortbildungen und deren Wirksamkeit aus der persönlichen Sicht von MusiklehrerInnen um die Rekonstruktion von Sinn- und Bedeutungszuweisungen handelt, ist die Arbeit forschungsmethodologisch als qualitative Untersuchung angelegt. Ausgangspunkt der Untersuchung sind acht leitfadengestützte, halbstandardisierte Interviews, die ich im Zeitraum von November 2005 bis August 2008 geführt habe. Die Auswertung erfolgte auf der Basis der Grounded Theory Methodologie (GTM)¹ nach Strauss (1994) bzw. Strauss/Corbin (1996). Dabei wurde das Kodierparadigma in Anlehnung an Grotluschen (2003) modifiziert. Nicht zuletzt auf Grund der eigenen beruflichen Nähe zum Forschungsfeld erschien es mir darüber hinaus wichtig, die Aspekte der von Breuer (2009) so genannten „Reflexiven Grounded Theory“ zu berücksichtigen.

Die interviewten Lehrerinnen und Lehrer wurden im Sinn des Theoretical Samplings im Verlauf des Auswertungsprozesses ausgewählt, wobei ich mich aus Gründen der Machbarkeit auf nordrhein-westfälische Schulen mit Sek. I- und Sek. II-Angebot beschränkt habe (also Gymnasien und Gesamtschulen, wobei einige der befragten Kolleginnen nicht in der Oberstufe unterrichten).

Der Forschungsgegenstand: Überblick über das Fortbildungsangebot im Untersuchungsraum

Die Frage nach dem Forschungsgegenstand scheint zunächst – ausgehend vom Forschungsinteresse – schnell beantwortet: Es geht um Fortbildungen für MusiklehrerInnen. Tatsächlich aber hat sich schon zu Beginn der Untersuchung als eine der zu klärenden Fragen herausgestellt, was eigentlich von den KollegInnen als Fortbildung angesehen bzw. auch, was als solche akzeptiert oder nicht akzeptiert wird, und wie die erkennbar unterschiedlichen Definitionen von „Fortbildung“ begründet werden. Damit ist neben der Erarbeitung der zentralen Forschungsfrage und der Schlüsselkategorien auch das Ausdifferenzieren des Forschungsgegenstandes selbst zum Teil der Analyse geworden.

1 Ich schließe mich der u. a. von Mey und Mruck (z. B. 2007: 11) empfohlenen und in der deutschsprachigen Literatur zunehmend üblichen Unterscheidung an, nach der als „Grounded Theory“ (GT) das Ergebnis einer entsprechenden empirischen Studie nach der „Grounded Theory Methodologie“ (GTM) bezeichnet wird.

In der offensichtlichsten Form werden unter Musik-Fortbildungen die Veranstaltungen „offizieller“ Anbieter verstanden, zu denen in erster Linie die Fortbildungen der Bezirksregierung gehören. So wird z. B. von der Bezirksregierung Köln im Fach Musik für Gymnasien in der Regel pro Halbjahr eine ganztägige Fortbildungsveranstaltung angeboten. Vorbereitet und organisiert werden die Veranstaltungen durch FachmoderatorInnen, also KollegInnen, die von der Bezirksregierung für diese Aufgabe gegen eine Unterrichtsentlastung abgestellt werden.² Inhaltlich werden diese Fortbildungen von den Interviewten so beschrieben, dass zumeist KollegInnen „gelungene“ Unterrichtsreihen aus der eigenen Praxis vorstellen, die zuvor von der Moderatorengruppe in einer Art Peer-Review-Verfahren ausgewählt, diskutiert und weiterentwickelt worden sind. Teile aus dieser Unterrichtsreihe werden in der Veranstaltung mit den TeilnehmerInnen praktisch erprobt. Ergänzend wird umfangreiches Material bereitgestellt, das neben Unterrichtsmaterial in Form von Arbeitsblättern und Sachinformationen auch die in der Fortbildung selbst bereits angesprochenen didaktischen und methodischen Überlegungen vertieft, wie z. B. die Frage nach der Einordnung in Bezug auf Lernbereiche des Lehrplans, Fragen nach möglichen Binnendifferenzierungen, manchmal auch Beispielklausuren etc. Diese Fortbildungen der Bezirksregierung sind häufig auf Oberstufenunterricht ausgerichtet. Dabei liegt in der Regel der Unterrichtsschwerpunkt auf der Reflexion; es können durchaus auch Praxisanteile vorkommen, allerdings nicht so vordergründig oder gar ausschließlich wie dies von den Interviewten etwa bei der Beschreibung von Fortbildungen der musikpädagogischen Verbände von den Interviewten genannt wird (s. u.).

Die zweite Form von Fortbildungen, die sich in den Interviews deutlich hervorhebt, sind Veranstaltungen der Verbände AfS und VDS, wobei der VDS als Anbieter im Zusammenhang mit der alle zwei Jahre veranstalteten überregionalen Bundesschulmusikwoche erwähnt wird, der AfS sowohl mit dem ebenfalls alle zwei Jahre veranstalteten Bundeskongress für Musikpädagogik als auch mit regelmäßig angebotenen regionalen Fortbildungen am Wochenende. Die genannten Fortbildungen werden beschrieben als überwiegend praxisorientierte Veranstaltungen: häufig instrumentalpraktisch ausgerichtet (zum Klassenmusizieren in unterschiedlichsten Varianten: Trommelkurse, Stomp, Boomwhacker-Workshops, Musizieren mit Rockinstrumenten, Gamelan-

2 Seit Anfang 2008 sind die ModeratorInnen Mitglieder so genannter „Kompetenzteams“ mit einer geänderten bzw. neuen Organisationsstruktur, in der u. a. auch schulform- und fachübergreifend gearbeitet wird; mehr dazu unter <http://www.kompetenzteams.nrw.de> (URL geprüft am 15. August 2009).

Musik, ...), aber auch zu Bereichen wie Tanz oder Schattentheater. Insbesondere die regionalen Angebote des AfS werden überwiegend in diesem Sinn beschrieben; neben dem eigenen Musizieren werden als Schwerpunkt dieser Fortbildungen methodische Überlegungen zur Vermittlung im Unterricht genannt, dagegen spielen Überlegungen auf der Meta-Ebene zu Fragen der Einbindung in die Inhalte eines allgemein bildenden Musikunterrichts, zur Anknüpfung an Bereiche des Lehrplans oder zu dahinter stehenden musikdidaktischen Konzeptionen in den Äußerungen der Interviewten in diesem Zusammenhang – im Unterschied zu den zuerst genannten Fortbildungen der Bezirksregierung – so gut wie keine Rolle.

Auch die Hochschule für Musik und Tanz Köln wird in den Interviews mit einer Fortbildungsveranstaltung erwähnt. Zwei der Interviewten haben an einem Kurs mit der Bezeichnung „Accompagnato – Brücken zwischen Studium und Beruf von Musiklehrenden“ teilgenommen, in dem Studierende mit ReferendarInnen und MusiklehrerInnen zusammengeführt werden. Das Kurskonzept ist damit gleichermaßen als Teil der studentischen Lehre wie als Fortbildungsveranstaltung angelegt. Die Veranstaltung erstreckt sich über ein Jahr: Neben acht über das Jahr verteilten dreistündigen Seminarterminen gibt es je einen ganztägigen Termin pro Semester, zusätzlich finden gegenseitige Unterrichtsbesuche in wechselnden Kleingruppen statt.³

Neben Bezirksregierung, musikpädagogischen Verbänden und Musikhochschule gibt es im Kölner Raum zahlreiche außerschulische Anbieter mit schulbezogenen Fortbildungsveranstaltungen; von den Interviewten genannt werden u. a. die Kölner Oper mit theaterpädagogischen Fortbildungen zur Szenischen Interpretation (jeweils angebunden an aktuelle Produktionen), Philharmonie und WDR mit Einführungen zu (Schüler-)Konzerten, die JazzHausSchule mit HipHop-Workshops und die Akademie Remscheid mit Veranstaltungen zu Theaterpädagogik. Ergänzt wird die Liste der genannten Angebote um solche, die nicht speziell auf Schule ausgerichtet sind, aber fortbildend wirksam geworden sind (z. B. Workshops auf der Frankfurter Musikmesse u. Ä.).

Aus den Äußerungen geht hervor, dass im Allgemeinen die Information über Angebote eher zufällig und nur unregelmäßig stattfindet. Natürlich lässt sich aus dem vorliegenden Datenmaterial auf Grund der Anlage als qualitative Studie mit einem theoriegeleiteten Sampling keine quantitative Aussage ableiten, aber man kann es zumindest als auffällig bezeichnen, dass zum einen das

3 vgl. Stöger 2004; Wiki Musikpädagogik 2010

gezielte Einholen von Informationen zu Fortbildungsangeboten bei den Interviewten die Ausnahme darstellt und zum anderen, dass insbesondere von den engagierten, vielfältig interessierten KollegInnen, die häufig an Fortbildungen auch unterschiedlicher Anbieter teilnehmen, beklagt wurde, dass es keine zentrale Informationsstelle für Fortbildungsangebote gibt, die mit einem Überblick über das existierende Angebot ein gezieltes, eigenverantwortlich organisiertes, systematisches Lernen im Beruf erleichtern würde.

Anhand der Beschreibungen des Angebots durch die Interviewten zeigt sich auch, dass sich Fortbildungen im Fach Musik sowohl vom Angebot her als auch in Bezug auf die Interessen der LehrerInnen grundsätzlich von Fortbildungen in anderen Fächern unterscheiden, weil sich in der unterschiedlichen Art der Veranstaltungen auch bereits sehr unterschiedliche Auffassungen von Musikunterricht und seinen Zielen dokumentieren können.

Im Zusammenhang mit dem Forschungsgegenstand war darüber hinaus interessant, dass bereits im ersten Interview der Begriff „Fortbildung“ sehr weit gefasst wurde. Herr A begann auf die eröffnende, sehr allgemein gehaltene Erzählaufforderung nach „grundsätzlichen Erfahrungen im Zusammenhang mit Fortbildungen“ mit der Aussage: „Ja, da muss man erst mal definieren, was eigentlich mit Fortbildung gemeint ist.“ Im weiteren Verlauf des Interviews bestätigte sich die methodische Erfahrung, dass häufig die erste Aussage eine Schlüsselfunktion für die Analyse des gesamten Interviews einnimmt: In diesem Fall gibt der erste Satz bereits einen Hinweis auf die vom Interviewten als paradox erfahrene Situation, dass er „Fortbildungen“ als (beruflich) nichtfortbildend erlebt hat, andere Lernangebote, die er ursprünglich eher „für sich selbst“ wahrgenommen hat, sich dagegen auf seinen Unterricht sehr positiv und nachhaltig ausgewirkt haben und somit für ihn beruflich fortbildend wurden, weil sie in Form einer Verbesserung seines Unterrichts wirksam geworden sind. Er äußert selbst die Auffassung, dass dies insbesondere für das Fach Musik typisch sein dürfte, weil sich hier private Interessen und berufliche Anforderungen häufig überschneiden.

Entwicklung des Kategoriensystems: die Schlüsselkategorien

Das o. g. erste forschungsleitende Interesse ist zu unterscheiden von den tatsächlichen Forschungsfragen, die schließlich die Auswertung bestimmten: Diese Forschungsfragen sind im Einklang mit der GTM erst im Lauf der Untersuchung konkretisiert und verfeinert worden; „Forschungsfragen“ kann man

begrifflich gewissermaßen als Gegenpol verstehen zu den aus dem Material entwickelten (Schlüssel-)kategorien, der bzw. den Kernkategorien und der damit verbundenen „Grounded Theory“, die diese beschreiben.

Eine wesentliche Schlüsselkategorie bei den Äußerungen zu Fortbildungen ist die der **Bewertung**. Dabei handelt es sich zumeist um rückblickende Äußerungen, in denen es um die Bewertung von Fortbildungen geht, an denen man teilgenommen hat (rückblickend, sofern nicht von den Interviewten generalisierend auf Fortbildungen Bezug genommen wird oder aufgrund von vorherigen Erfahrungen antizipierende Bewertungen vorgenommen werden). Eine zweite Schlüsselkategorie ist dagegen in der Regel zukunftsgerichtet, sie wurde bereits bei der Beschreibung des Forschungsgegenstands angesprochen und stimmt mit dem Ausgangsinteresse überein: die Kategorie der **Wünsche**, die von den Interviewten konkret in Bezug auf Fortbildungen geäußert werden. Diese Wünsche beziehen sich häufig auf die Funktion „Lernen“ mit dem Ziel der Unterrichtswirksamkeit; dabei geht es zum einen um inhaltliche Aspekte (Themen, Gegenstände), zum anderen aber auch um Aspekte wie methodische Anregungen oder den Wunsch nach einer kontinuierlichen Begleitung längerer Lernprozesse. Bei der Auswertung hat sich als weniger aussagekräftig herausgestellt, was konkret in Form einzelner Themen oder Gegenstände gewünscht wird als vielmehr, in welchem Kontext und mit welchen Begründungen diese Wünsche geäußert werden – also Fragen, denen in der GTM nach Strauss/Corbin im Rahmen des axialen Kodierens mit Hilfe des Kodierparadigmas als heuristisch-analytischem Rahmen nachgegangen wird.⁴ Wünsche werden dabei zunächst vor allem im Hinblick auf dasjenige geäußert, was als ein zentrales Ziel von Fortbildungen angesehen wird: Sich Fortbilden verstanden als Lernen im Hinblick auf den Beruf. Betont wird aber auch die Bedeutung des Lernens „für sich selbst“ als einem Lernen, das nicht

4 Während es beim offenen Kodieren darum geht, weitgehend unvoreingenommen, ohne vorgefasste theoretische Annahmen und in enger Anlehnung an das empirische Material möglichst viele thematisch relevante Konzepte herauszuarbeiten, kommen beim axialen Kodieren bereits stärker Auswahlentscheidungen zum Tragen, indem bestimmte, beim offenen Kodieren gewonnene Konzepte als übergeordnete bzw. für die sich allmählich abzeichnende Theorie besonders bedeutsame Konzepte in den Mittelpunkt der Analyse gestellt werden vgl. Strübing 2008). Dabei dient das von Strauss/Corbin vorgeschlagene Kodierparadigma einer Systematisierung der Analyse, indem mit Hilfe eines heuristischen Rahmens „um die Achse“ eines Konzepts bzw. einer übergeordneten Kategorie herum kodiert wird (vgl. Strauss/Corbin 1996).

unmittelbar unterrichtswirksam werden muss (aber u. U. indirekt etwa durch den eigenen Wissenszuwachs wirksam werden kann). Erkennbar werden also an dieser Stelle auch Wünsche, die nicht ausschließlich berufsbezogen sind; Letzteres gilt auch für andere Wünsche an Fortbildungen, die nicht unmittelbar auf die Kernfunktion „Lernen“ bezogen sind.

Neben diesen konkret benannten lassen sich weitere Wünsche aus jenen Äußerungen ableiten, die darauf hinweisen, welche Faktoren zur Teilnahme an einer konkreten Fortbildung motivieren bzw. welche ein Hindernis für die Teilnahme darstellen können. Diese Kategorie der geäußerten **Motive zur Teilnahme** an konkreten Veranstaltungen sowie die Kategorie „Wünsche“ überschneiden sich inhaltlich zum Teil, die genannten Motive zur Teilnahme erweitern das Gesamtbild jedoch auch. Ein Beispiel dafür ist das Ziel eines Vermerks der Teilnahme in der Personalakte, weil der Interviewte sich demnächst auf eine Beförderungsstelle bewerben möchte und sein berufliches Engagement dokumentieren will, wofür er institutionalisierte Fortbildungsmaßnahmen benötigt; ein weiteres Beispiel ist das Interesse am Austausch mit FachkollegInnen, weil man sich an der Schule mit dem Fach Musik isoliert fühlt.

Auch im Zusammenhang mit der Kategorie „Bewertungen“ werden Erwartungen (oder zumindest Hoffnungen) in Bezug auf Fortbildungen erkennbar, ohne dass diese explizit als Wünsche geäußert würden. Neben dieser Ergänzung stehen die Kategorien „Wünsche“ und „Bewertungen“ aber natürlich auch dergestalt in Beziehung zueinander, dass Bewertungen z. T. davon abhängig sind, ob Wünsche erfüllt werden oder nicht. Es zeigte sich, dass zum einen Wünsche nicht alle erfüllt sein müssen, damit eine Fortbildung positiv bewertet wird, sondern dass es zum Teil Einzelkriterien gibt, deren alleinige Erfüllung für eine positive Bewertung bereits ausreicht; zum anderen gibt es aber auch (je nach Person unterschiedliche) Ausschlusskriterien für eine positive Bewertung. Interessant ist in diesem Zusammenhang, dass das Lernen im Hinblick auf das eigene Unterrichten, obwohl es entsprechend der geäußerten Wünsche als zentrales Element von Fortbildungen angesehen wird, dennoch nicht zu den notwendigen Bedingungen für eine positiven Bewertung gehört.

Die genannten drei Kategorien ergänzen sich also, indem sie aus verschiedenen Perspektiven ein Bild der **Bedürfnisse**⁵ vermitteln, die MusiklehrerInnen im Zusammenhang mit Fortbildungen erkennen lassen. Als zweite zentrale Schlüsselkategorie wird in den Äußerungen die der **Wirksamkeit** von Fortbildungen deutlich, dabei insbesondere die Unterrichtswirksamkeit, die wiederum unter verschiedenen Aspekten ausdifferenziert werden kann.

Bedürfnisse von MusiklehrerInnen im Zusammenhang mit Fortbildungen

Der Blick der interviewten MusiklehrerInnen auf Fortbildungen wird v. a. durch ihr Interesse an unterrichtsbezogenem Lernen⁶ bestimmt. Neben konkreten Aussagen diesbezüglich wird diese Sichtweise auch mehrfach indirekt bestärkt, z. B. durch die Äußerung, wonach der „private“ Jazzpiano-Unterricht eigentlich eine „Fortbildung“ darstellt, da er unterrichtswirksam wird. Umgekehrt werden aber, wie oben bereits erwähnt, Fortbildungen nicht generell als schlecht bewertet, wenn sie *nicht* unterrichtswirksam werden: Unterrichtswirksamkeit ist also einerseits das wichtigste Definitionskriterium für Fortbildungen, ihr Fehlen andererseits aber kein Ausschlusskriterium für eine positive Bewertung.

Diese Diskrepanz ist zum einen erkennbar das Resultat entsprechender Erfahrungen mit Fortbildungen: Solche Veranstaltungen werden von einigen Interviewten als häufig bzw. sogar meistens nicht unterrichtswirksam und trotzdem als positive Erfahrung beschrieben. Zum anderen aber wird dadurch in den Äußerungen der LehrerInnen deutlich, dass auch unabhängig von der unmittelbaren Unterrichtswirksamkeit Bedürfnisse existieren, die durch Fortbildungen abgedeckt werden und dann für eine positive Bewertung sorgen können. Es lassen sich also Bedürfnisse unterscheiden, die unmittelbar mit einer

5 Die Kategorie „Wünsche“ umfasst konkrete, explizit von den Interviewten benannte Aspekte, von denen als abstrahierende Kategorie die dahinter liegenden, zusammenfassenden bzw. übergeordneten „Bedürfnisse“ unterschieden werden.

6 Der Begriff „Lernen“ wird im Rahmen dieses Textes – so wie auch in den Interviews erkennbar – zunächst einmal in der Bedeutung eines Zugewinns an Wissen, Fertigkeiten und Fähigkeiten verwendet, ohne ihn in den Kontext einer bestimmten Lerntheorie zu stellen. Vor diesem Hintergrund ist dann im Hinblick auf die Fragestellung von besonderem Interesse, was alles im Sinn eines professionellen Wissens in den Äußerungen erfasst wird und inwiefern sich dadurch ein unterschiedlicher Blick auf Fortbildungen und ihre Wirksamkeit ergibt.

Wirksamkeit im Unterricht verbunden sind von solchen, die nicht oder nur indirekt auf eine Unterrichtswirksamkeit abzielen.

Bedürfnisse unabhängig von unmittelbarer Unterrichtswirksamkeit

Eine erste Erklärung für positive Bewertungen trotz fehlender Unterrichtswirksamkeit ist, dass auch im Zusammenhang mit institutionalisierten Fortbildungen das Lernen selbst als eigenständiges Bedürfnis erkennbar wird, das **Lernen um des Lernens willen**. Es geht dabei um eine Erweiterung des eigenen Wissens im Sinn einer persönlichen Entwicklung ohne die Zweckgebundenheit einer (unmittelbaren) Unterrichtswirksamkeit, weil die jeweiligen Inhalte über das für den Unterricht verwendbare Wissen hinaus gehen. So sagt beispielsweise Herr A: „Auch wenn man die niemals machen würde ... wenn ich einfach mal was zum Beispiel über die Leningrad-Sinfonie von Schostakovich gelernt hab, ist das schon 'ne gelungene Fortbildung. Wenn ich einfach mal wieder selber was gelernt hab. Und vielleicht sehr gelungen ist, wenn ich es im Unterricht anwende oder erfolgreich anwende, aber (.) das muss gar nicht unbedingt sein.“⁷

Auch im Kontext dieses Lernens ohne eine direkte Anwendungsmöglichkeit im Unterricht wird dennoch immer wieder ein Bezug hergestellt, indem das Lernen als **Motivation** für das eigene Unterrichten und damit als eine Grundlage für mehr Berufszufriedenheit dargestellt wird. Im Gegensatz zu sonst überwiegenden Äußerungen, wenn es um unmittelbare Unterrichtswirksamkeit geht, werden dann auch Veranstaltungen als positiv bewertet, die auf eine rezeptive Teilnahme hin konzipiert sind. In Bezug auf einen Bundeskongress des AfS sagt Frau B: „So'n Vorlesungscharakter gab's ja auch, also (.) das fand ich für mich auch wiederum gut, weil, das Studium ist ja jetzt doch wieder 'ne Zeit her, und als Oberstufenlehrer find' ich das auch nicht schlecht, dass man ab und zu mal wieder intellektuelle Anregungen kriegt.“ Der Aspekt einer verbesserten Berufsmotivation durch die Teilnahme an Fortbildungen wird auch im Zusammenhang mit musikpraktischen Angeboten erkennbar: Auch wenn die Fortbildungsinhalte z. B. mangels Ausstattung an Instrumenten oder vermuteter Überforderung der eigenen SchülerInnen als nicht im Unter-

7 Anmerkung zu den hier verwendeten vereinfachten Transkriptionsregeln: Die Unterstreichung steht für eine auffällige Betonung, (2) steht für eine Sprechpause von etwa zwei Sekunden und (.) für Pausen, die weniger als eine Sekunde dauern.

nicht umsetzbar beschrieben werden, können sie dennoch die Motivation des eigenen Unterrichtens mit musikpraktischen Anteilen erhöhen.

Unabhängig vom konkreten Unterrichtsbezug ist also das grundsätzliche Interesse an der eigenen, **vielfältigen Auseinandersetzung mit Musik**, sei es kognitiv oder musikpraktisch, ein wichtiger Grund für die Teilnahme an Fortbildungen. In diesem Zusammenhang ist in den Interviews häufig die Rede vom „Spaß“ als Kriterium für eine gute Fortbildung. Der **Erlebnisaspekt** selbst stellt damit ein Bedürfnis dar. Dabei erscheint die Wahl des Begriffs „Spaß“ auf den ersten Blick möglicherweise anrühlich, dessen tatsächliche Bedeutung wird aber durch den dargestellten Kontext deutlich: Die bei oberflächlicher Betrachtung leicht missverständliche Wortwahl wird im Zusammenhang mit Lernen und mit Musizieren verwendet und gleichzeitig in den Kontext zu Motivation und Berufszufriedenheit gestellt; damit unterstreichen die interviewten LehrerInnen die grundsätzliche Bedeutung von Fortbildungen für ihre Motivation. Gleichzeitig zeigt sich aber auch eine deutliche Parallele zu den eigenen Unterrichtszielen; differenziert wird u. a. in einer Aussage von Frau H erkennbar, was sie mit „Spaß“ verbindet und wie eng dies mit mehreren ihrer Unterrichtsziele zusammenhängt: „Diese (2) ja, Freude am praktischen Musizieren, dass das einen bereichert, ... dass man irgendwie zusammen arbeitet, oder wenn Schüler so nach 'nem Probenwochenende sagen »boah das hat so viel Spaß gemacht« und wenn man dann mal aufzählt die Stunden, die wir wirklich geprobt haben und wie wenig Freizeit, und trotzdem solche Schüler dann sagen »das war toll«, ... wenn so (..) man den Schülern beigebracht kriegt, dass sich da so'n bisschen Mühe lohnt, ... und dieses gemeinsame Musizieren ja nun ein Gemeinschaftserlebnis bringt, und man das einfach gut fand, ... so Spass, merken, dass Musik einen innerlich bereichern kann, wenn man das so rüber bringen könnte, ist denke ich schon viel.“

Weitere Bedürfnisse ohne unmittelbaren Unterrichtsbezug werden auch über das Lernen hinaus genannt. Für alle im Rahmen dieser Darstellung einzeln ausgewiesenen Bedürfnisse ist aber festzuhalten, dass sie als untereinander vernetzt zu verstehen sind, wodurch sich letztlich auch immer wieder ein Unterrichtsbezug abzeichnet. Dazu gehört auch das immer wieder betonte Bedürfnis nach einem **Austausch mit FachkollegInnen**. Dieser ist sowohl inhaltsbezogen als auch (berufs- und fachbezogen) emotional gemeint. Ziel ist es, damit aus einer empfundenen Isoliertheit bei der Unterrichtsarbeit und in der Auseinandersetzung mit beruflichen Schwierigkeiten herauszukommen. Durch den Kontext der Schilderungen werden auch hier wieder Bezüge zum unterrichtsbezogenen eigenen Lernen deutlich, z. B. wenn in der weiterfüh-

renden Begründung für den Wunsch nach fachlichem Austausch positiv erlebte Beispiele herangezogen werden: Der Austausch mit ReferendarInnen und Studierenden, die Ideen aus den Studienseminaren oder Hochschulen mitbringen, oder der Austausch mit FachleiterInnen im Rahmen von gelungenen Unterrichtsnachbesprechungen werden als Lernerfahrungen im Sinn eines kollegialen Austauschs jenseits von institutionalisierten Fortbildungen beschrieben, die im eigenen Unterricht wirksam geworden sind. Frau F bezeichnet diese Art des fachlichen Austauschs als eine „Kultur der Teamarbeit“, die sie durch Fortbildungen gefördert sehen möchte. Selbst an Schulen mit mehreren MusiklehrerInnen kommt Teamarbeit unter KollegInnen nach den vorliegenden Aussagen nur ausnahmsweise (z. B. wegen fächerübergreifender Kurse) zustande; eine Ausnahme hierbei stellen in den Interviews die KollegInnen dar, die an Schulen arbeiten, in denen Musik einen eigenen Schwerpunkt im Schulprogramm darstellt und Teil des Schulentwicklungsprozesses ist. Ansonsten findet Teamarbeit im Sinn von gemeinsamem berufsbezogenem Lernen aber scheinbar eher zufällig und nicht geplant statt, selbst wenn der Austausch mit FachkollegInnen in dieser Hinsicht als positiv angesehen wird: Es fehlt die Verankerung der Idee eines gemeinsamen Lernens im beruflichen Alltag.

Eng mit dem Wunsch nach Austausch verbunden ist das Bedürfnis nach **Partizipation** im Kontext von Fortbildungen, das in verschiedenen Formen erkennbar wird. Im einfachsten Fall geht es darum, die eigenen Wünsche in die Planung von Fortbildungsveranstaltungen einbringen zu können. Darüber hinaus besteht aber auch ein Interesse an der Mitarbeit bei der Entwicklung und Durchführung von Fortbildungen. Eine Begründung dafür ist, dass gerade dies als wirkungsvollste Form des eigenen berufsbezogenen Lernens angesehen wird, also gewissermaßen eine „Fortbildung durch Fortbilden“ in Anlehnung an die Idee des „Lernens durch Lehren“. Besonders deutlich wird dies bei einer Fachmoderatorin, die zu den interviewten Lehrerinnen im Rahmen dieser Befragung gehörte. Sie bezeichnete das oben beschriebene gemeinsame Diskutieren und Vorbereiten von Fortbildungen in der Moderatorengruppe als die für ihren eigenen Unterricht wirksamste „Fortbildung“. Eine weitere erkennbare Begründung für das Bedürfnis nach Partizipation stellt das Bedürfnis nach **Anerkennung und Bestätigung** durch KollegInnen dar, also nach einer Wertschätzung aus professioneller Sicht in Bezug auf die eigene unterrichtliche Planungs- und Entwicklungsarbeit, weil dies scheinbar in dieser Form im Unterrichtsalltag sonst kaum erfahren wird. Nicht zuletzt wird auch in der Rolle als TeilnehmerIn das Bedürfnis nach Partizipation dergestalt deutlich, dass Veranstaltungen ausgesprochen negativ bewertet werden, in denen Referen-

tInnen mit den Worten von Herrn A als „Oberlehrer“ auftreten und die TeilnehmerInnen sich in eine (im negativen Sinn hierarchisierend gemeinte) Schülerrolle gezwungen sehen, wie dies Frau B beschreibt: „Dann denke ich auch, derjenige hat nicht die Stärke, jetzt gleichberechtigte Partner neben sich zu haben, sondern er will uns da irgendwo 'ne Ebene tiefer halten.“

Bedürfnisse im Hinblick auf unmittelbare Unterrichtswirksamkeit

Wenn in den Interviews ausdrücklich das **unterrichtsbezogene Lernen** als Bedürfnis thematisiert wird, dann wird dies auf zwei Arten begründet: zum einen schlicht mit dem dahinter liegenden Bedürfnis nach Entlastung durch Hilfestellung und Unterstützung bei der alltäglichen Arbeit der Unterrichtsvorbereitung, zum anderen mit dem Bedürfnis nach einer persönlichen Weiterentwicklung und damit dem Ziel der eigenen Professionalisierung.

Das Bedürfnis nach **Entlastung** bezieht sich v. a. auf Sachwissen, dahinter steht der Wunsch nach Lernen durch Information (z. B. der Vermittlung von Analysen der Werke aus den Vorgaben des Zentralabiturs) und der Wunsch nach entsprechendem Unterrichtsmaterial (z. B. didaktisch reduzierte Partituren oder ausgewählte, aufbereitete musikwissenschaftliche Materialien und ausgewählte Hörbeispiele). Im Kontext des Bedürfnisses nach einer Reduzierung der Arbeitsbelastung wird „Fortbildung“ also eher als Teamwork im Sinn von Arbeitsteilung verstanden, wobei mehrfach geäußert wird, dass i. d. R. die entsprechenden Materialien auch ohne die Veranstaltung selbst ihren Zweck erfüllen würden. Entlastung durch Arbeitsteilung ist nach Meinung der erwähnten Fachmoderatorin auch eines der wichtigsten Ziele von Fortbildungen der Bezirksregierung.

Über den Aspekt der Entlastung hinaus geht das Bedürfnis nach einer **persönlichen Weiterentwicklung**, also nach Lernen bezogen auf ein umfassend zu verstehendes professionelles Wissen im Hinblick auf die individuelle **Unterrichtsentwicklung**. Mit Unterrichtsentwicklung ist hier in Abgrenzung von der Unterrichtsplanung konkret das Verändern des eigenen Unterrichtens gemeint. Dies kann sich auf Inhalte und auf Methoden beziehen; in beiden Fällen wird als Motiv einerseits erkennbar, den Unterricht zu erweitern oder zu erneuern sowie empfundene eigene Defizite aufzuarbeiten (inhaltlich z. B. wegen Unsicherheit beim Unterrichten von Pflichtthemen für das Zentralabitur, aber auch instrumentalpraktisch z. B. auf Grund von Ausbildungslücken im Bereich Populärer Musik) oder andererseits Unterrichtsprobleme methodischer Art zu beheben.

Zu solchen Problemen gehört der Umgang mit Unterrichtsstörungen beim Klassenmusizieren oder die Schwierigkeit, in Bezug auf bestimmte Unterrichtsinhalte die SchülerInnen nicht zu erreichen. Herr A wünscht sich etwa beim Thema „Formenlehre“ methodische Hilfe dazu, wie man die SchülerInnen erreichen und z. B. für bestimmte Themen des Zentralabiturs interessieren kann: „Hier diese A-Dur-Sinfonie ... da hätt ich zum Beispiel mal gern ‚ne Fortbildung, wie man zum Beispiel ‚ne pfiffige Unterrichtsreihe über Formenlehre machen kann.“ Auf besonders dramatische Weise schildert Herr C, dass sich für ihn die grundsätzliche Schwierigkeit, die SchülerInnen zu erreichen, erst im Lauf der Zeit entwickelt und dann so weit verstärkt hat, dass inzwischen für ihn das Ziel des heutigen Musikunterrichts selbst Anlass für Verunsicherung ist: Er unterrichtet seit etwas mehr als dreißig Jahren Musik an einem Gymnasium und vermisst heute die anfangs noch empfundene Sicherheit hinsichtlich der Übereinstimmung eigener Vorstellungen von Musikunterricht, veränderten (und sich weiter verändernden) Vorgaben und sich verändernden Schülerinteressen. Herr C wünscht sich deshalb von Fortbildungen Antworten auf die Frage nach dem „Sinn“ von Musikunterricht: „Ich würd' sogar noch einen draufsetzen, sogar Sinnfragen nachgehen. Wenn wir Musik unterrichten, was unterrichten wir da eigentlich? (2) Ist heute nicht mehr klar. Als ich anfing, war das glockenklar, aber (.) ich frag' mich das (.) in mancher Klasse wirklich (.) unter hohem Druck stehend immer wieder, ... um was geht's da eigentlich.“

Die inhaltlichen Wünsche im Hinblick auf die persönliche Weiterentwicklung werden (neben dem auch hier wieder erkennbaren Wunsch der Entlastung beim Einarbeiten) vor allem damit begründet, dass die gewünschten unbekannten Bereiche als unzugänglich erscheinen – es geht also zunächst um ein Bedürfnis nach **Zugangswissen** im Sinn eines ersten Zugangs zu einem Wissensbereich. So beschreibt z. B. Frau E, die kurz vor der Pensionierung steht und eine typische kunstmusikorientierte Ausbildung erfahren hat, eine Orientierungslosigkeit in Bereichen der Populären Musik. Sie wünscht sich „solides Wissen“, um einzelne Musiktitel oder Bands bestimmten Stilen zuordnen zu können. Herr A dagegen, der zum Zeitpunkt des Interviews seit zehn Jahren am Gymnasium unterrichtete, hat zwar auch (mangels Alternativen) ein kunstmusikorientiertes Lehramtsstudium abgeschlossen, sieht sich von seinen musikalischen Präferenzen her aber eher als „Jazzler“ und hat keinen Bedarf an Fortbildungen im Bereich Populärer Musik, weil er in diesem Bereich schon viel in seinem Unterricht anbietet. Er fühlt sich zwar in Bezug auf die traditionelle europäische Kunstmusik einigermaßen gut ausgebildet, weil sein Studi-

um im Wesentlichen darauf ausgerichtet war, Orientierungshilfe wünscht er aber dennoch im Zusammenhang mit den Werkvorgaben bei der Einführung des Zentralabiturs.⁸ Das Zentralabitur wird auch in anderen Interviews in diesem Kontext thematisiert, wenn die bisher unbekannten Gegenstände und die Frage nach möglichen Quellen zur Aneignung des erforderlichen eigenen Wissens für eine Behandlung dieser Gegenstände im Unterricht für Unsicherheit sorgen; mit Zugangswissen kann also – neben einer Vermittlung im Rahmen der Veranstaltung selbst – auch **Quellenwissen** gemeint sein. Betont wird die Bedeutung von Fortbildungen gerade im Zusammenhang mit Pflichtthemen beim Zentralabitur auch von Frau E: Diese Fortbildungen werden aus ihrer Sicht von den KollegInnen gerne angenommen, „wenn Sie's selber nicht schaffen ... sich was anzuknobeln, manche können's ja auch, aber eben nicht alle.“ Außerdem verbirgt sich dahinter aber die zusätzliche Problematik in NRW, dass die Einführung des Zentralabiturs im Fach Musik selbst, ungeachtet der Inhalte, für Verunsicherung v. a. bei jüngeren KollegInnen gesorgt hat. Zum einen waren für sie bei der Einführung die unterrichtlichen Zielsetzungen im Zusammenhang mit den geforderten Werken nicht im Detail erkennbar, zum anderen wurde die Idee eines Zentralabiturs an sich als widersprüchlich zum gleichzeitig gültigen Lehrplan für die Sekundarstufe II empfunden, der auf die Partizipation der SchülerInnen bei der Unterrichtsgestaltung Wert legt und diese ausdrücklich einfordert: Das Zentralabitur legt Gegenstände verbindlich fest, während der Lehrplan das „Verhandeln“ von Themen und Gegenständen (im Rahmen der Obligatorik) mit den SchülerInnen verlangt.⁹

Ein weiterer Bereich, der im Zusammenhang mit Zugangswissen als Bedürfnis erkennbar wird, ist die Musik fremder Kulturen, wie z. B. indische Musik. Mehrfach ausdrücklich davon ausgenommen wird hierbei Musik aus Schwarzafrika, da hier die Materiallage von den Interviewten als gut eingeschätzt wird und sie sich gut informiert fühlen, v. a. durch das musikpädagogische Standardwerk von Volker Schütz (1992), aber auch durch viele Querverweise in der musikpädagogischen Literatur und in Unterrichtswerken im Kontext von afroamerikanischer Populärer Musik und ihren Wurzeln. Auch im interkulturellen Zusammenhang wird als Begründung für Themenwünsche an-

8 Das Zentralabitur wurde in NRW im Jahr 2007 eingeführt.

9 Mit beiden Schwierigkeiten hatten die interviewten älteren KollegInnen tendenziell weniger Probleme, da sie nach ihren Aussagen zumindest teilweise auf frühere Unterrichtsreihen und Konzeptionen und auf musikdidaktische Vorstellungen früherer Lehrpläne zurückgreifen konnten.

gegeben, dass die grundlegenden Orientierungsmöglichkeiten fehlen: der persönliche Zugang zur Musik und wenigstens ansatzweise zu einem Verständnis der Kultur, aber auch zum Musizieren und damit zu Möglichkeiten der angemessenen musikpraktischen Auseinandersetzung im Unterricht. Neben den oben bereits genannten Wissensaspekten spielen also bei dem Bedürfnis nach Zugangswissen auch instrumentalpraktische Fähigkeiten eine Rolle, die in diesem Zusammenhang als wichtig erscheinen. Über das gegenstandsbezogene Wissen im engeren Sinn hinaus wünscht sich Frau B beim Themenfeld „Musik fremder Kulturen“ Hilfe bei der Orientierung über die ortsansässige Musikszene fremder Kulturen und möglichst auch die Vermittlung von Kontakten zu MusikerInnen oder Veranstaltern, um darauf im Unterricht oder in Projektarbeit und im Rahmen von Exkursionen oder Veranstaltungsbesuchen zurückgreifen zu können; auch das genannte Quellenwissen ist also in umfassendem Sinn zu verstehen.

Neben inhaltlichen Wissensgebieten bezieht sich der Wunsch nach Unterrichtsentwicklung auch auf Zugangswissen in Bezug auf (musikpraktische) Fähigkeiten und Fertigkeiten, wobei die zusätzliche Funktion genannt wird, dass Fortbildungen häufig überhaupt erst den erforderlichen **Mut machen**, sich mit bestimmten, bisher unzugänglich erscheinenden Bereichen (wie z. B. Improvisation, Percussion, szenische Interpretation) auseinanderzusetzen und ein Einbeziehen in die eigene Unterrichtsentwicklung in Betracht zu ziehen. Auf der anderen Seite existiert auch der Wunsch nach einer Vertiefung von Fähigkeiten v. a. im instrumentalbereich und in diesem Zusammenhang nach dem Begleiten von längerfristigen Übeprozessen in solchen Bereichen, in denen bereits ein gewisser „Expertenstatus“ im Hinblick auf das professionelle Wissen erreicht worden ist; Letzteres wird im Zusammenhang mit Fortbildungen aber eher als Utopie denn als realistischer Wunsch genannt.

Unterrichtswirksamkeit von Fortbildungen

Im Zusammenhang mit den hier im Überblick dargestellten Bedürfnissen ist bereits deutlich geworden, dass die Bedeutung von Fortbildungen für MusiklehrerInnen auch durch die **Wirksamkeit** bestimmt wird, die ihnen von den Interviewten zugesprochen wird.

Die Wirksamkeit ist, wie bereits angesprochen, nicht auf den unmittelbaren Unterricht beschränkt, sondern kann sich auch indirekt oder zu einem späteren Zeitpunkt auf Unterricht oder allgemeiner auf Schule bzw. auf das Berufsleben

über den eigentlichen Unterricht hinaus beziehen; auch Wirkungen jenseits des Berufslebens werden genannt. Wichtig ist den interviewten MusiklehrerInnen aber – teilweise trotz fehlender positiver Erfahrungen diesbezüglich – insbesondere der Aspekt der Unterrichtswirksamkeit von Fortbildungen. Dies wird nicht zuletzt dadurch deutlich, dass auch bei nicht unmittelbar unterrichtsbezogenen Bedürfnissen immer wieder ein Bezug zum Unterricht bzw. zum Unterrichten hergestellt wird.

Von besonderem Interesse hinsichtlich der Wirksamkeit ist die Frage, welches Verhältnis sich aus Sicht der Befragten zwischen den geschilderten Bedürfnissen und der Teilnahme an Veranstaltungen ergibt, also z. B. inwieweit Fortbildungen zu einer Entlastung beitragen oder weitergehend auch Unterrichtsentwicklung ermöglichen. Zwei Extremfälle zeichnen sich diesbezüglich ab: Von einigen Befragten wird keinerlei unmittelbare Wirksamkeit der erlebten Fortbildungen für den eigenen Unterricht gesehen; begründet wird dies in den Interviews u. a. mit der fehlenden Relevanz der angebotenen Inhalte für die eigenen Schülergruppen auf Grund eines zu anspruchsvollen bzw. zu abstrakten Niveaus oder mit fehlender Relevanz auf Grund der Beliebigkeit der gewählten Gegenstände. Bei einer Interviewpartnerin wurden Fortbildungen aber im Kontrast zur völlig fehlenden Unterrichtswirksamkeit sogar zur entscheidenden Instanz für die Ausgestaltung des Unterrichts: Frau H beschreibt, dass der eigene Unterricht zu Anfang des Berufslebens „viel zu kopflastig“ gewesen sei und kaum Praxisanteile enthalten habe, weil in der Ausbildung keine Grundlagen für einen praxisorientierten Unterricht gelegt worden seien. Der Unterricht habe sich insbesondere durch den Besuch von Fortbildungen verändert: „Also ich glaube, ohne die Fortbildungen wäre ich viel frustrierter und würde vielleicht noch Dinge machen, die ich vor zehn Jahren gemacht hatte.“ Dabei hat sich aber offensichtlich nicht nur der Praxisanteil des Unterrichts erhöht, sondern das gesamte Verständnis von gelingendem Musikunterricht hat sich verändert. Diese Änderung sieht konkret so aus, dass aus heutiger Sicht von Frau H die Themen „Musiktheorie“ und „Analyse“ im Unterricht mit *ihren* SchülerInnen gar keinen Sinn und ihnen auch keinen Spaß machen würden, woraus dann wiederum persönlicher Frust für sie als Lehrerin resultiere. „Theorie“ unterrichtet sie nach eigener Aussage deshalb eher „zum Disziplinieren“ und Kräftesparen, „weil man nicht jede Stunde Praxis machen kann“. In diesem Fall wird das über die Jahre wahrgenommene Fortbildungsangebot als entscheidende Instanz dafür genannt, wie der eigene Musikunterricht inzwischen aussieht, während Lehrplan, schulinternes Curriculum, Studi-

um oder Referendariat für diesen Unterricht inzwischen nur noch eine sehr untergeordnete Rolle spielen.

Zwischen diesen beiden Polen bewegen sich die Äußerungen der übrigen MusiklehrerInnen, die grundsätzlich eher nicht erwarten, dass Fortbildungen unmittelbar wirksam werden; ihre Erwartungen sind offensichtlich auf Grund der erlebten Fortbildungen im Lauf der Zeit reduziert worden. Dies betrifft vor allem die Vorstellung, Material oder vorgestellte Unterrichtsreihen eins-zu-eins übernehmen zu können; davon haben die Interviewten während der ersten Jahre ihres Berufslebens Abstand genommen. Frau B sagt z. B.: „Man kann nicht 'ne Pauschalreihe entwerfen, das geht sowieso nicht. Jeder hat seine eigene Klientel, und seine Schüler vor sich und seine Materialien und seine spezielle Art zu unterrichten.“ Vielfältige Möglichkeiten werden dagegen hinsichtlich der Wirksamkeit im Sinn eines „Steinbruchs“ gesehen, sei es in Bezug auf das Material, die Methoden, vermittelte Kompetenzen oder auch nur Ideen, die man mitgenommen hat: „Und dann isses auch bei mir so, dass ich wirklich vieles lange aufhebe, wie so'n Schwamm alles aufsaugt und irgendwann braucht man's dann oder findet 'ne Idee, es mit was anderem zu vernetzen. Also ich denke, das muss bei mir jetzt gar nicht so unmittelbar jetzt irgendwie Ergebnisse bringen.“ (Herr F).

Ausblick: Bedingungen für Unterrichtswirksamkeit

Zusammenfassend kann bis hierhin festgehalten werden, dass die verschiedenen Bedürfnisse im Kontext unterrichtsbezogenen Lernens sich zum einen unterschiedlichen Wissensarten im Sinn eines umfassenden Begriffs von professionellem Wissen zuordnen lassen; zum anderen lässt sich erkennen, dass das Lernen bei Fortbildungen im Hinblick auf Unterrichtsentwicklung als Teil eines Prozesses zu verstehen ist, da sich die geäußerten Bedürfnisse auf unterschiedliche Phasen eines Wissenserwerbs beziehen: vom „Zugangswissen“ (im Sinn von Einstiegs- und Quellenwissen), auf das ein eigenständiges oder begleitetes vertiefendes Lernen aufbauen kann, bis hin zu einem „Expertenwissen“ in Bezug auf die Umsetzung im Unterricht. Dieses Verständnis verschiedener Phasen zeigt sich auch in Wünschen nach dem Aufzeigen von Zugängen zu Wissensbereichen oder zu Lernwegen einerseits und in Wünschen nach Anleitung und Begleitung von Lernprozessen durch längerfristige Veranstaltungen andererseits. Diese Phasen sind vor allem wichtig in Bezug auf die Frage, unter welchen Bedingungen Fortbildungen tatsächlich im Unterricht und bei der Unterrichtsentwicklung wirksam werden.

Auf die Bedingungen für eine Unterrichtswirksamkeit von Fortbildungen kann im Rahmen dieses Beitrags nur noch kurz eingegangen werden, ausführlich wird dies in der Veröffentlichung der Studie erfolgen: Ob eine Fortbildungserfahrung in den Unterricht eingebracht wird, ist grundsätzlich abhängig von der Entscheidung der jeweiligen Lehrperson, diese Erfahrung in die eigene Unterrichtsplanung oder -entwicklung einzubeziehen. Bei dieser Entscheidung spielen offenbar mehrere Faktoren zusammen, die sich drei **Bedingungsfeldern** zuordnen lassen, nämlich

1. die persönliche Entwicklung im Hinblick auf das professionelle Wissen (hierbei ist v. a. die Einschätzung der TeilnehmerInnen über das erreichte oder erreichbare Gelingen des eigenen Entwicklungsprozesses im Sinn der oben angesprochenen Phasen des Wissenserwerbs bedeutsam),
2. die Anschlussmöglichkeit an Unterrichtsziele und den persönlichen Unterrichtsstil (Unterrichtsziele können vorgegeben oder selbst gewählt sein, stehen also in Abhängigkeit von Vorgaben des Zentralabiturs, von Richtlinien, schulinternem Lehrplan und Schulprogramm, von Schülerinteressen bzw. deren Möglichkeiten und von den eigenen Einstellungen; der Unterrichtsstil bezieht sich z. B. darauf, ob Klassenmusizieren grundsätzlich als zu anstrengend empfunden wird o. Ä.) und
3. die Verfügbarkeit materieller Ressourcen (Unterrichtsmaterial, Medien/Instrumente, in erweitertem Verständnis auch: zeitliche Ressourcen).

Wenn die Bedingungen in allen drei Feldern als positiv (im Sinn von entweder in ausreichendem Maße gegeben oder realistisch erreichbar) eingeschätzt werden, sprechen die befragten LehrerInnen Fortbildungen die Möglichkeit zu, ihre eigene Unterrichtsentwicklung zu beeinflussen und in ihrem Unterricht wirksam zu werden. Vor dem Hintergrund dieser Bedingungsfelder beantwortet sich also im Hinblick auf Fortbildungen jeweils individuell die Frage, ob eine Unterrichtsentwicklung ermöglicht worden ist. Das Wirksamwerden von Fortbildungen im Unterricht ist dabei selbst wiederum als Prozess zu verstehen, den Niessen (2006) im Individualkonzept von MusiklehrerInnen bereits als Lernprozess beschrieben hat, was sich auch in den hier ausgewerteten Interviews bestätigt.

Fortbildungen als 3. Phase der Lehrerbildung?

Eine ausführliche Diskussion der Ergebnisse hinsichtlich möglicher Konsequenzen kann an dieser Stelle nicht erfolgen. Es sollen aber einige Fragen an-

gesprochen werden, die sich m. E. angesichts der Sichtweise der befragten LehrerInnen auf Fortbildungen geradezu aufdrängen: Von allen Befragten wird die Notwendigkeit von Fortbildung gesehen.¹⁰ Begründet wird dies u. a. mit (als unvermeidbar angesehenen) Ausbildungslücken und vielfältigen Veränderungen im Lauf des Berufslebens. Wenn Fortbildungen aber in diesem Sinn als 3. Phase der Lehrerbildung „funktionieren“ sollen, wenn sie also bezogen auf Unterricht nicht nur der Entlastung dienen, sondern auch zur Professionalisierung von LehrerInnen beitragen sollen, kann die Diskrepanz zwischen positiver Beurteilung von Fortbildungen durch die LehrerInnen bei gleichzeitig (mehr oder weniger hingenommener) fehlender Wirksamkeit im Hinblick auf die persönliche Unterrichtsentwicklung nicht zufrieden stellen.

Angesichts der Komplexität der hier nur in Schwerpunkten dargestellten Bedürfnisse und Bedingungen für Unterrichtswirksamkeit erscheint es dabei wenig hilfreich, die häufig beklagte mangelnde Wirksamkeit von Fortbildungen im Hinblick auf Unterrichtsentwicklung pauschal einer „Fortbildungsabstinenz“ der LehrerInnen zuzuschreiben, hinter der Bastian/Hafen (1990: 73) „eine gehörige Portion Bequemlichkeit, Nachlässigkeit, mangelndes Fortbildungsbewußtsein und berufsbedingte (?) Inflexibilität“ vermuten.¹¹

Vielmehr lassen die hier in einem ersten Überblick vorgestellten Ergebnisse danach fragen, wie Fortbildung angesichts des Missverhältnisses zwischen unterrichtsbezogenen Bedürfnissen einerseits und tatsächlich empfundener Unterrichtswirksamkeit andererseits zu einer Professionalisierung im Rahmen der 3. Phase der Lehrerbildung beitragen könnten. Damit steht weniger die Frage nach bestimmten Inhalten im Mittelpunkt; die Vielfalt der genannten Wünsche und deren Heterogenität lässt zwar durch ihre Begründungen sinnvolle Schwerpunkte für Fortbildungsthemen erkennen, die auch im Interesse eines bedarfs- und bedürfnisdeckenden Angebots abgefragt werden sollten. Die konkreten Gegenstände einer Veranstaltung können aber dennoch immer nur exemplarisch sein, und für eine mögliche Integration in das eigene Unterrichten wird deshalb auf die Fortbildungsinhalte bezogen unter dem Stichwort „Steinbruch“ gewünscht, „alltagstaugliche“ Bausteine anzubieten.

10 Dies deckt sich mit ebenso deutlichen Ergebnissen einer quantitative Pilotstudie von Bastian/Hafen (1990).

11 Kritisiert wurde diese Zuweisung bereits von Niessen (2006: 154) im Zusammenhang mit dem Individualkonzept von MusiklehrerInnen.

Grundsätzlich stellt sich aber die weitergehende Frage danach, inwieweit das aktuelle Angebot an Fortbildungen durch alternative Veranstaltungsformen ergänzt werden könnte, die auch die bisher vernachlässigten unter den genannten Bedürfnissen berücksichtigen und eine Unterrichtswirksamkeit vor dem Hintergrund der Bedingungsfelder ermöglichen würden. Der oben beschriebene, von zwei der befragten LehrerInnen besuchte Accompagnato-Kurs greift einige solcher Bedürfnisse auf, die in anderen Fortbildungen selten eine Rolle spielen (v. a. den Aspekt der Partizipation und der längerfristigen Begleitung von Lernprozessen); der Kurs wird allerdings eher als allgemeinpädagogische Veranstaltung angesehen und hilft durch diese andere Zielsetzung nur begrenzt im Kontext von spezifisch musikpädagogischen Fragestellungen bzw. bei der konkreten persönlichen Unterrichtsentwicklung im Fach Musik. Ein weiterer Fortbildungsansatz, in dem die Partizipation der TeilnehmerInnen u. a. in Form von Aktionsforschung bewusst in eine Fortbildung einbezogen wurde und der dabei ausdrücklich auf das Fach Musik ausgerichtet war, findet sich in der Literatur bei Eibach/Knolle/Münch (2006) in der Beschreibung des Me[i]Mus-Modellvorhabens. Als Teil eines BLK-Programms hatte diese Veranstaltung aber einen singulären Charakter und ist als Veranstaltungsform nicht in der regulären Fortbildungslandschaft verankert.

Letztlich stellt sich auch zum einen die Frage, inwieweit ein organisatorischer Rahmen für solche neuen Fortbildungsformen bereitgestellt werden kann und zum anderen die Frage einer institutionellen Anbindung musikpädagogischer Fortbildungen, denn es sind zwar viele Fragen der „Lehrerbildung im Beruf“ von fächerübergreifender Natur (wie sie mancherorts auch von Zentren für Lehrerbildung u. ä. genannten Einrichtungen aufgegriffen werden oder aufgegriffen werden sollten), aber in den Interviews wird deutlich, dass es nach Meinung der befragten LehrerInnen insbesondere für die Musikpädagogik fachspezifische Anforderungen an Fortbildungen im Sinn einer notwendigen 3. Phase der Lehrerbildung gibt.

Literatur

Bastian, Hans Günther; Hafen, Roland (1990): (Musik-)Lehrer über ihre Fortbildung oder Lehrer sind fortbildungseinsichtig, aber fortbildungsabstinent. In: Ott, Thomas; Günther, Ulrich (Hg.): Lehrerbildung und Lehrerweiterbildung für den Musikunterricht, (Gegenwartsfragen der Musikpädagogik. Schriftenreihe der Bundesfachgruppe Musikpädagogik, Bd. 3), Essen: Die Blaue Eule, S. 26–77.

- Breuer, Franz (2009): *Reflexive Grounded Theory. Eine Einführung für die Forschungspraxis*. Wiesbaden: VS Verlag.
- Eibach, Martin; Münch, Thomas; Knolle, Niels (2006): MusiklehrerInnen als Lernende – Überlegungen zu Strategien der Fortbildung im Fach Musik. In: Knolle, Niels (Hg.): *Lehr- und Lernforschung in der Musikpädagogik*, (Musikpädagogische Forschung 27), Essen: Die Blaue Eule, S. 43–69.
- Grotluschen, Anke (2003): *Widerständiges Lernen im Web – virtuell selbstbestimmt? Eine qualitative Studie über E-Learning in der beruflichen Erwachsenenbildung*. Münster [u.a.]: Waxmann.
- Mey, Günter; Mruck, Katja (2007): *Grounded Theory Methodologie – Bemerkungen zu einem prominenten Forschungsstil*. In: Mey, Günter; Mruck, Katja (Hg.): *Grounded Theory Reader. (Historical Social Research, Supplement No. 19)*, Köln: Zentrum für Historische Sozialforschung, S. 11–39.
- Niessen, Anne (2006): *Individualkonzepte von Musiklehrern*. Berlin [u.a.]: LIT.
- Schütz, Volker (1992): *Musik in Schwarzafrika. Arbeitsbuch für den Musikunterricht in den Sekundarstufen*. Oldershausen: Institut für populäre Musik.
- Stöger, Christine (2004): Erfahrungswissen für den Lehrberuf. „*accompagnato*“ – Ein Kurskonzept zur Vernetzung von Studium und Beruf. In: *Diskussion Musikpädagogik*, 23/2004, S. 33–38.
- Strauss, Anselm L. (1994): *Grundlagen Qualitativer Sozialforschung: Datenanalyse und Theoriebildung in der empirischen und soziologischen Forschung* (= Uni-Taschenbücher, Bd. 1776). München: W. Fink Verlag.
- Strauss, Anselm L.; Corbin, Juliet (1996, Orig. 1990): *Grounded Theory: Grundlagen Qualitativer Sozialforschung*. Weinheim: Beltz/Psychologie Verlags Union.
- Strübing, Jörg (2008): *Grounded Theory. Zur sozialtheoretischen und epistemologischen Fundierung des Verfahrens der empirisch begründeten Theoriebildung*. 2. überarbeitete und erweiterte Auflage. (= *Qualitative Sozialforschung*, 15). Wiesbaden: VS Verlag.
- Wiki Musikpädagogik (20.02.2010): *Accompagnato*. URL: <http://www.dirk-bechtel.de/wiki/index.php/Accompagnato> (Zugriff: 20.02.2010)

Lampenfieber und Aufführungsängste bei Kindern und Jugendlichen

Erhebungen zur Selbstwahrnehmung im Rahmen musikalischer Vortragssituationen

Fragestellung

Kinder und Jugendliche standen im Zusammenhang mit musikalischen Vortragssituationen im Gegensatz zu Erwachsenen bislang kaum im Fokus der empirischen Lampenfieberforschung. Zwar wurden in einer Studie des Deutschen Musikrates (1992)¹ Jugendliche im Alter zwischen 14 und 21 Jahren untersucht, dennoch handelte es sich wiederum um eine spezielle Subgruppe, nämlich Teilnehmer des Wettbewerbs „Jugend musiziert“. Die Studie deckte aber dennoch interessante Zusammenhänge auf. Auf der Grundlage der Erfassung von Art und Stärke des Zusammenhangs zwischen Leistung und Lampenfieber konnte u.a. ein positiv signifikanter, aber gleichwohl geringer Zusammenhang zwischen „Leistung“ und „Lampenfieber“ festgestellt werden, d.h., dass höheres Lampenfieber demnach schlechtere Leistung bedeutet. Ferner wurden in dieser Studie auch Unterschiede zwischen Mädchen und Jungen hinsichtlich der Publikumsangst untersucht. Insgesamt gaben dabei 59,9 % der Mädchen im Vergleich zu nur 40,5 % der Jungen an, unter Lampenfieber zu leiden.

Aufgrund der wenigen erhobenen Daten bei Kindern und Jugendlichen setzte sich die vorliegende Studie zum Ziel, Erkenntnisse über die Ausprägungen des Phänomens bei dieser Altersgruppe zu erlangen.

Unter Aufführungsangst wird in dieser Studie eine spezielle Form der Publikumsangst verstanden, die durch Aspekte der sozialen Angst und der Leis-

1 vgl. Linzenkirchner, P. (1995): Lampenfieber und musikalische Leistung, In: Gembris, Heiner u.a. (Hrsg.): Musikpädagogische Forschungsberichte, Augsburg: Wißner, S.73-85

tungsangst beeinflusst wird und als eine Zustandsangst beschrieben werden kann. Aufführungsangst ist somit abhängig von der individuellen Angstneigung einer Person. Diese Angst kann sich sowohl in emotional-physiologischen Symptomen zeigen, als auch in Form von Besorgnis- und Angstgedanken auf der Ebene der Kognitionen äußern. Vom Begriff Aufführungsangst ist dagegen der Begriff Lampenfieber als anregendes und leistungsförderndes Phänomen abzugrenzen.

Auf der Grundlage dieser Definitionen von Aufführungsangst und Lampenfieber fanden die Aspekte Kognition, Emotion und Wahrnehmung (psycho)somatischer Begleitsymptome in der Erhebung Berücksichtigung. Im Speziellen sollte dabei untersucht werden, ob sich bezüglich der Aufführungsangst Unterschiede in verschiedenen Subgruppen, das Alter und das Geschlecht betreffend, ergeben und ob das Phänomen „Lampenfieber“ von den Kindern und Jugendlichen sowohl als hemmend als auch als fördernd wahrgenommen wird.

Durchführung und Methode

Für die quantitativ ausgerichtete Erhebung (durchgeführt in 2008) konnte aus insgesamt vier Musikschulen des Rems-Murr-Kreises und des angrenzenden Ostalbkreises eine Stichprobe von 265 InstrumentalschülerInnen realisiert werden (insgesamt 800 ausgegebene Fragebögen, Rücklaufquote 34,4%). Den 175 weiblichen und 90 männlichen InstrumentalschülerInnen im Alter von 8-19 Jahren wurde dabei ein standardisierter Fragebogen in Anlehnung an das TAI-G (vgl. Hodapp 1991) vorgelegt. Aufgrund der teilweise sehr komplexen Fragestellungen war eine Eingrenzung bezüglich des Alters notwendig. So erwies sich in einer Pretest-Erhebung der konstruierte Fragebogen und die Problematik des Themas für Kinder unter 8 Jahren als zu schwierig und komplex.

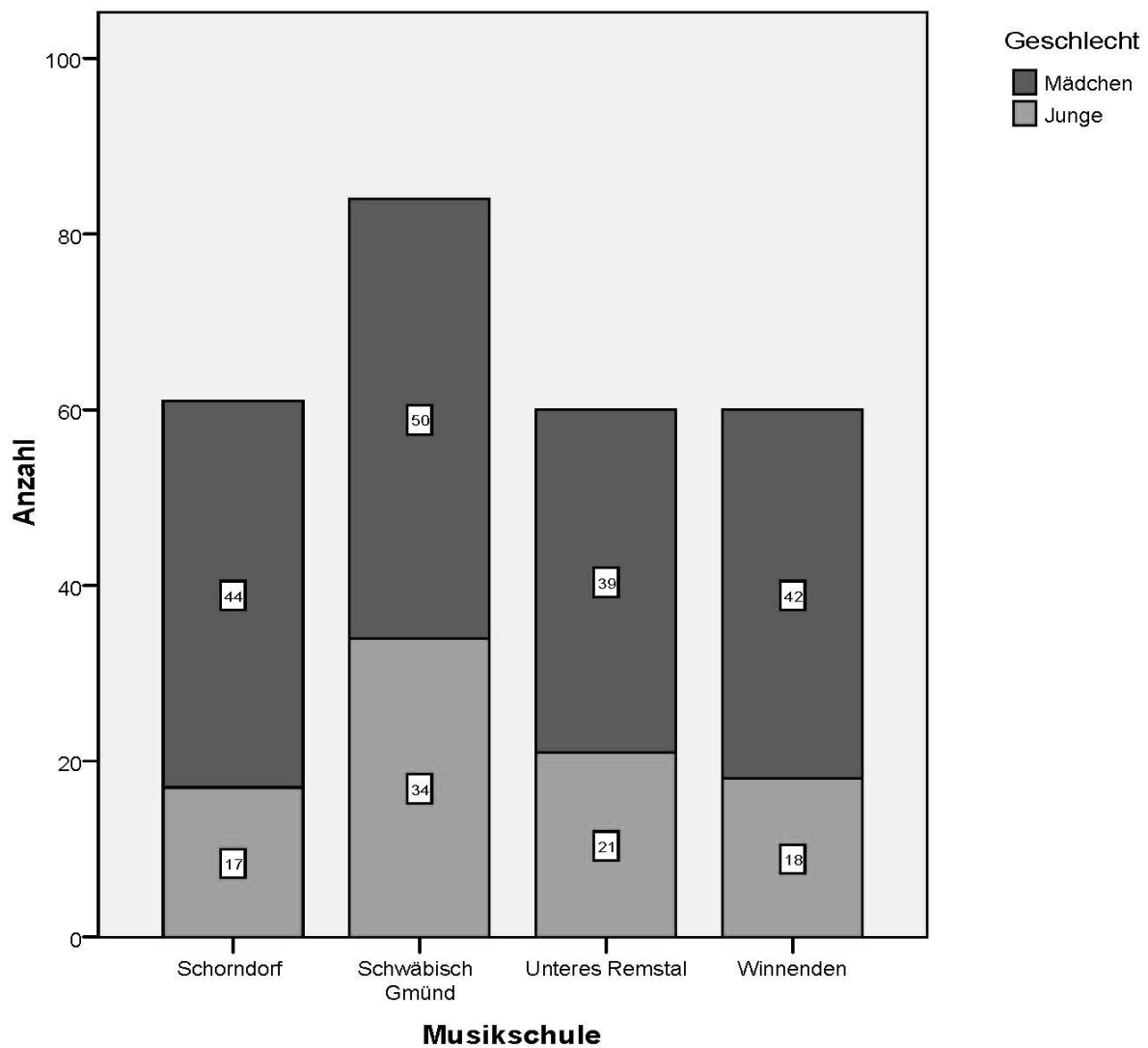


Abb. 1: Stichprobe

Erfragt wurden zum einen die subjektive Selbstwahrnehmung für die beiden Vorstellungssituationen „vor musikalischen Vorträgen“ bzw. „direkt während des Auftritts“ über Items der Variablen „Aufgeregtheit“ und „Besorgtheit“.

Die Items für die Variablen wurden teilweise dem TAI-G entnommen und modifiziert sowie, inspiriert durch eine Reihe anderer Ängstlichkeitstests und -fragebogen, durch weitere Items auf die Summe von neun Items pro Komponente ergänzt. Zur Verhinderung von Response Sets, (hier: Akquieszenzen), wurden dabei jeweils zwei Items umgekehrt bzw. umgepolt.

Aufgeregtheit:

- „Ich spüre ein komisches Gefühl (Kribbeln) im Bauch.“
- „Ich bin am ganzen Körper verkrampft.“
- „Das Herz schlägt mir bis zum Hals.“
- „Ich habe ein beklemmendes Gefühl.“
- „Ich bin aufgeregt.“
- „Ich bin entspannt.“ (umgekehrtes Item)
- „Meine Hände und oder meine Knie zittern.“
- „Ich schwitze am ganzen Körper.“
- „Ich bin ruhig und gelassen (locker).“ (umgekehrtes Item)

Besorgtheit:

- „Ich mache mir Sorgen, ob ich auch alles schaffe.“
- „Ich frage mich, ob meine Leistung ausreicht.“
- „Ich denke daran, wie wichtig mir ein gutes Ergebnis ist.“
- „Ich bin besorgt, dass etwas schief laufen könnte.“
- „Ich habe Angst etwas falsch zu machen.“
- „Ich weiß, dass ich gut bin und mich gut vorbereitet habe.“ (umgekehrtes Item)
- „Ich möchte die Erwartungen meiner Eltern/Lehrer erfüllen.“
- „Ich bin sicher, dass ich alles kann.“ (umgekehrtes Item)
- „Ich hoffe, dass es bald vorbei ist.“

Alle Items wurden mit Hilfe von Rating-Skalen erfasst. Die Aufteilung der Rating-Skalen erfolgte dabei nach verbalen Marken mittels einer vierfach gestuften Likert-Skala mit folgenden Antwortkategorien: fast immer – oft – manchmal – fast nie. Zur Vermeidung einer zentralen Tendenz wurde auf die Wahl einer neutralen Kategorie (z. B. gelegentlich) verzichtet. Diese insgesamt 18 Items wurden durch ein Item zur „Allgemeinen Aufregung“ ergänzt. Für dieses Item wurde in Anlehnung an die „Visual Analogue Scale“ für

ängstliche Kinder (Bernstein/Garfinkel 1992) ein symbolisches Rating eigens für diese Erhebung entwickelt.

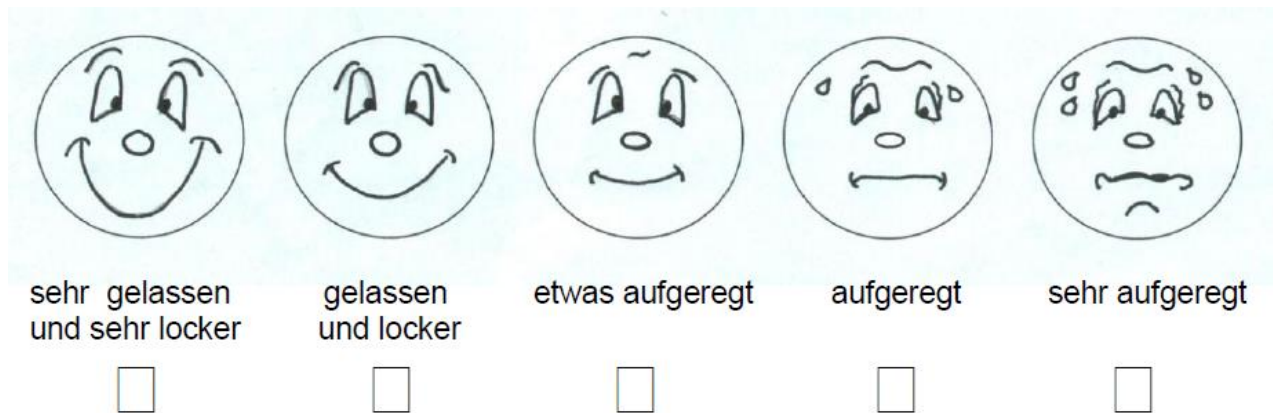


Abb. 2: Symbolische Rating-Skala

Zudem erfolgte eine Befragung zur positiven bzw. negativen Wahrnehmung des Phänomens. Diese erfolgte ebenfalls durch den Einsatz von Rating-Skalen, welche durch drei offene Fragestellungen ergänzt wurden: Was ist für dich Lampenfieber? Schildere, was mit dir kurz vor einer Vortragssituation passiert. Denkst du, dass dir Lampenfieber beim Vorspielen hilft? Warum?

Statistische Analysen und Ergebnisse

Zur Überprüfung der Geschlechterunterschiede hinsichtlich der Konstrukte „Besorgtheit“ und „Aufgeregtheit“ wurden Mittelwertsvergleiche mithilfe von t-Tests nach Student durchgeführt. Zur Kontrolle der Ergebnisse wurden zusätzlich die einzelnen Items mittels eines nichtparametrischen Verfahrens, dem sogenannten U-Test von Mann und Whitney, in den Blick genommen. Nach einem Mittelwertsvergleich für die beiden Vorstellungssituationen ist mit einer zweiseitigen Signifikanz von $\alpha = .000$ für die Variablen „Aufgeregtheit“ und „Besorgtheit“ davon auszugehen, dass sich die beiden Geschlechtergruppen hinsichtlich der beiden Variablen höchst signifikant voneinander unterscheiden. Dieses Messergebnis konnte mit der Ausnahme von 3 Items („*Ich möchte die Erwartungen meiner Eltern/Lehrer erfüllen.*“/ „*Ich denke daran, wie wichtig mir ein gutes Ergebnis ist.*“/ „*Ich schwitze am ganzen Körper.*“) durch das nichtparametrische Verfahren bestätigt werden. Es ist hier jedoch anzumerken, dass dieses Ergebnis noch einmal näher zu untersuchen ist, denn es kann nach dieser Erhebung noch nicht als gesichert gelten, dass Mädchen tatsächlich häufiger Aufführungsangst wahrnehmen als Jungen. Es wäre ebenso denkbar, dass

das Messergebnis darauf zurückzuführen ist, dass Mädchen lediglich eher dazu bereit sind Angaben zu dieser Problematik zu machen.

Zur Untersuchung der altersbedingten Unterschiede erfolgte zunächst eine Kategorisierung der MusikschülerInnen in drei verschiedene Altersgruppen (8-11 Jahre, 12-15 Jahre, 16-19 Jahre). Die statistische Auswertung erfolgte über dieselben Analyseinstrumente. Dabei konnte festgestellt werden, dass es eine signifikante Wahrnehmungsveränderung von Aufführungsangst bezüglich der Gruppierungsvariablen „Alter“ gibt. Je älter der/die Vorspielende ist, desto bewusster und intensiver nimmt er bzw. sie Aufführungsangst wahr. Bei der „Aufgeregtheits“-Variable war dabei ein früherer Zeitpunkt der Veränderung auszumachen als bei der „Besorgtheits“-Variable.

Nach einer Korrelationsanalyse nach Pearson konnte für die beiden verschiedenen Vorstellungssituationen ein hoher ($r = .739$ / vor dem musikalischen Vortrag) und ein mittlerer ($r = .681$ / während des musikalischen Vortrags) gleichläufiger Zusammenhang zwischen den Variablen „Aufgeregtheit“ und „Besorgtheit“ festgehalten werden.

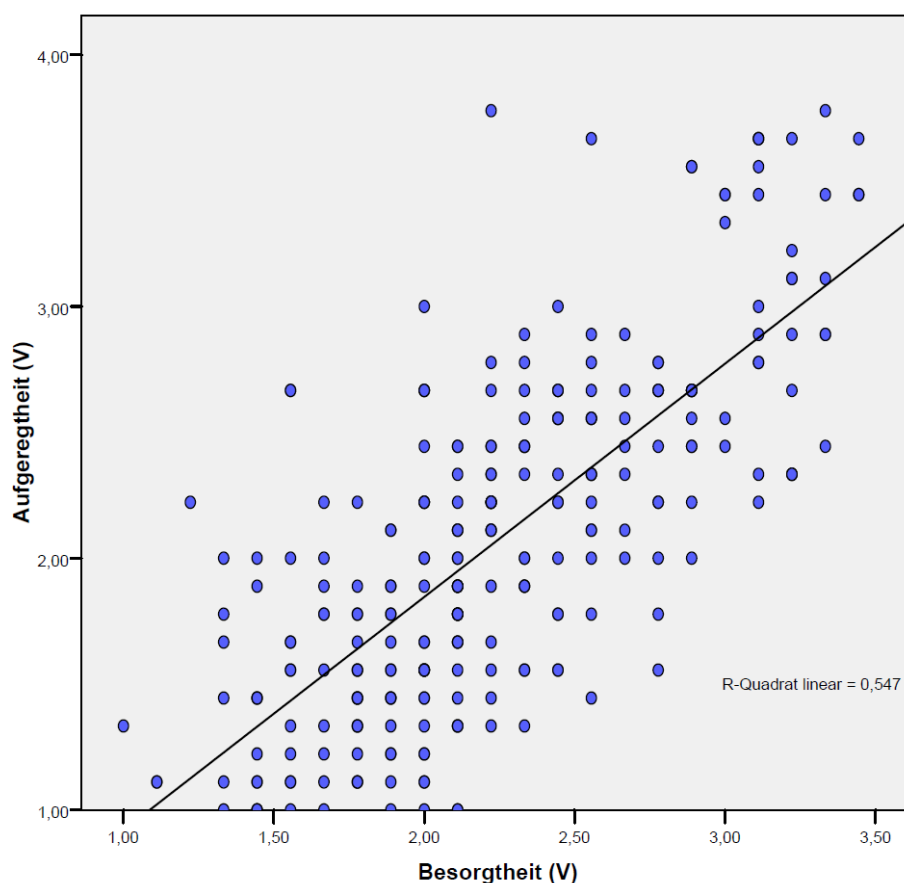


Abb. 3: Korrelation „Vor dem Auftritt“

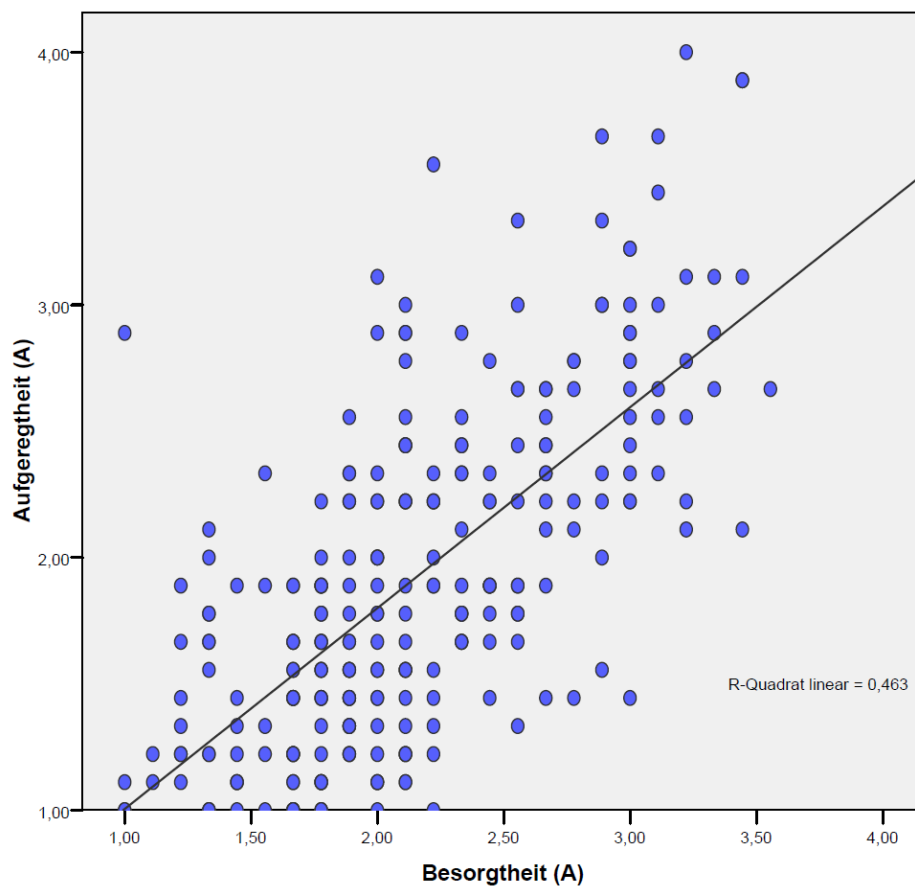


Abb. 4: Korrelation „Während des Auftritts“

Über einen t-Test für verbundene Stichproben konnte zudem festgestellt werden, dass die Wahrnehmung von Aufführungsangst vor dem Auftritt signifikant höher ist als während des musikalischen Vortrags ($\alpha = .000$). Zur Überprüfung der Veränderungshypothese über die einzelnen Itempaare wurde auf einen sogenannten Wilcoxon-Test zurückgegriffen. Mithilfe dieses Tests konnten folgende Itempaare als nicht mitbestimmende Faktoren ausgeschlossen werden: „*Meine Hände und/oder meine Knie zittern.*“ und „*Ich schwitze am ganzen Körper.*“, „*Ich bin sicher, dass ich alles kann.*“ und „*Ich weiß, dass ich gut bin und mich gut vorbereitet habe.*“ Diese genannten Items können somit als für die Veränderung unbedeutend gelten.

Neben diesen statistischen Analysen wurden die offenen Fragestellungen ausgewertet: Dabei kann festgehalten werden, dass für Kinder und Jugendliche sowohl Aufführungsängste als leistungshemmendes Phänomen existiert, als auch Lampenfieber, das als leistungsfördernd wahrgenommen wird. Ein Musikschüler (16 Jahre) beschreibt das Phänomen beispielsweise wie folgt:

„Das hat zwei Seiten: Man hat Respekt vor der Zuschauermenge und gibt sich Mühe, alles richtig zu machen. Manchmal ist man sehr aufgeregt und kann sich deshalb nicht richtig konzentrieren.“

Andere MusikschrülerInnen nehmen Lampenfieber hauptsächlich als leistungssteigernd wahr. So zum Beispiel ein 13-jähriges Mädchen:

„Ja, wenn ich aufgeregt bin und allen beweisen will, was ich kann, konzentriere ich mich mehr und kann somit höhere Leistung erbringen.“

Ebenso gibt es aber auch MusikschrülerInnen, die explizit auf eine hemmende Wirkung von „Lampenfieber“ hinweisen. Stellvertretend sei hier das klare Statement eines 12-jährigen Jungen genannt:

„Ich denke nicht, dass das Lampenfieber hilft, denn in der Aufregung macht man Fehler.“

Die Frage, ob Lampenfieber beim Vorspielen einen positiven Effekt haben kann, beantwortete eine 14-jährige Musikschrülerin folgendermaßen:

„Lampenfieber hilft vor allem dann, wenn man es während oder kurz vor dem Auftritt überwinden kann. So Lampenfieber gehört nun mal zu einem Auftritt dazu. Wenn man das Lampenfieber überwunden hat, ist man nach dem gelungenen Auftritt umso freier, gelassener und glücklicher.“

In eine ähnliche Richtung geht die Aussage eines 13-Jährigen Jungen:

„Nein, ich denke nicht, dass mir starkes Lampenfieber hilft, aber es ist sicher oft von Nutzen, wenn ich ein bisschen Lampenfieber habe.“

Insgesamt scheint Lampenfieber bzw. Aufführungsangst beim Vorspielen mit einem Instrument für Kinder und Jugendliche eine große Rolle zu spielen, ungeachtet dessen, ob dieses Phänomen als leistungshemmend oder -steigernd wahrgenommen wird.

Resümee

Die vorliegenden Erkenntnisse bestätigen die Relevanz von Aufführungsängsten bei Kindern und Jugendlichen. Weiterführende Untersuchungen sind insbesondere über die Bedingungsfaktoren von Aufführungsängsten in dieser Altersgruppe angezeigt. Auf dieser Basis wäre das Entwickeln gezielter präventiver Maßnahmen für den Instrumentalunterricht möglich, um Vorspielängste zu reduzieren und die beflügelnden Eigenschaften des Lampenfiebers nutzbar zu machen.

Literatur

- Bernstein, G. A./Garfinkel, B. D. (1992): The Visual Analogue Scale for Anxiety-Revised: Psychometric Properties. *Journal of Anxiety Disorders* 6, P. 223-239.
- Essau, C. A. (2003): Angst bei Kindern und Jugendlichen, München u.a.: Ernst Reinhardt Verlag.
- Federer, M./Schneider, S./Margraf, J./Herrle, J. (2000): Angstsensitivität und Angststörungen bei Achtjährigen. *Kindheit und Entwicklung* 9, H. 4, S. 241-250.
- Hodapp, V. (1991): Das Prüfungsängstlichkeitsinventar TAI-G: Eine erweiterte und modifizierte Version mit vier Komponenten. *Zeitschrift für Pädagogische Psychologie* 5, H. 2, S. 121-130.
- Hodapp, V./Laux, L./Spielberger, C.D. (1982): Theorie und Messung der emotionalen und kognitiven Komponente der Prüfungsangst. *Zeitschrift für Differentielle und Diagnostische Psychologie* 3, H. 3, S.169-184.
- Hoyer, J./ Margraf, J. (Hrsg.) (2003): Angstdiagnostik. Grundlagen und Testverfahren, Berlin u.a.: Springer.
- Krahwehl, I./Altenmüller, E. (2000): Lampenfieber unter Musikstudenten: Häufigkeit, Ausprägung und „heimliche Theorien“. *Musikphysiologie und Musikermedizin* 7, H. 4, S. 173-182.
- Krohne, H.W. (1996): Angst und Angstbewältigung, Stuttgart u.a.: Kohlhammer.
- Laux, L./Glanzmann, P./Schaffner, P./Spielberger, C.D. (1981): Das State-Trait-Angstinventar (STAI), Weinheim: Beltz.

- Linzenkirchner, P. (1993): Lampenfieber bei Teilnehmern des Wettbewerbs „Jugend musiziert“, In: Gembris, Heiner u.a. (Hrsg.): Musikpädagogische Forschungsberichte, Augsburg: Wißner, S. 267-278.
- Linzenkirchner, P. (1995): Lampenfieber und musikalische Leistung, In: Gembris, Heiner u.a. (Hrsg.): Musikpädagogische Forschungsberichte, Augsburg: Wißner, S.73-85.
- Linzenkirchner, P./Eger-Harsch, G. (1995): Gute Noten mit kritischen Anmerkungen. Wirkungsanalyse der Wettbewerbe „Jugend musiziert“ 1984 bis 1993. Dokumentation und Kommentierung, München: Deutscher Musikrat.
- Lohaus, A./Eschenbeck, H./Kohlmann, C.-W./Klein-Heßling, J. (2006): SSKJ 3-8 Fragebogen zur Erhebung von Stress und Stressbewältigung im Kindes- und Jugendalter, Göttingen u.a.: Hogrefe.
- Möller, H. (2002): Lampenfieber und Aufführungsängste sind nicht dasselbe!, In: Hofmann, G./ Trübsbach, C. (Hrsg.): Mensch & Musik – Diskussionsbeiträge im Schnittpunkt von Musik, Medizin, Physiologie und Psychologie, Augsburg: Wißner, S. 45-58.
- Motte-Haber, H. de la (2005): Die Musikerpersönlichkeit. Charakterzüge, Leistung und Lampenfieber, Selbstkonzept, In: Motte-Haber, H. de la/Rötter, G. (Hrsg.): Musikpsychologie, Laaber: Laaber-Verlag, S. 515-551.
- Musch, J./Bröder, A. (1999): Psychometrische Eigenschaften und Validität des multidimensionalen Prüfungsängstlichkeitsinventars TAI-G, Zeitschrift für Pädagogische Psychologie 13, H. 1/2, S. 100-105.
- Wieczerkowski, W./Nickel, H./Janowski, A./Fittkau, B./Rauer, W. (1974): Angstfragebogen für Schüler (AFS), Braunschweig: Westermann.

Qualität entsteht beim Lernen

Lerner integrierende Qualitätsbeurteilung beim E-Learning

Gliederung

1. Was ist Qualität - Zur Problematik eines multidimensionalen Konstruktes
2. Qualität im Bildungsbereich –
3. Paradigmenwechsel, Subjektorientierung und Dienstleistungsverhältnis
4. Lernerorientierung als Imperativ – Konsequenzen für die empirische Forschung
5. Lerner integrierende Qualitätsbeurteilung am Beispiel von Musik-Lernsoftware

1 Was ist Qualität - Zur Problematik eines multidimensionalen Konstruktes

Qualität ist ein schwer fassbares, fast schillerndes Konstrukt. Eine prägnante Begriffsbestimmung ist kaum möglich, „Qualität“ wird in zahlreichen Theorien und wissenschaftlichen und wirtschaftlichen Disziplinen unterschiedlich definiert, expliziert und exemplifiziert. Ein Geschäftsführer versteht etwas anderes darunter als ein Softwareentwickler, der Staat etwas anderes als ein Lehrer, der Nuklearmediziner etwas anderes als der Bio-Bauer. Institutionen zur Qualitätssicherung werden gegründet, im Wirtschaftssektor („Deutsche Gesellschaft für Qualität“¹) ebenso so wie im Bildungsbereich („Institut zur Qualitätsentwicklung im Bildungswesen“²). Gütesiegel werden eingeführt (Bio-Siegel, Blauer Engel, GS-Zeichen), die Stiftung Warentest untersucht und bewertet Waren und Dienstleistungen verschiedener Anbieter, die Testergebnisse

1 <http://www.dgq.de> (18.02.2010)

2 <http://www.iqb.hu-berlin.de> (18.02.2010)

bestimmen Verbrauchermeinungen und Verkaufszahlen.³ Die Fachliteratur weist auf die „*verheerende Wirkung der täglichen Werbung auf das richtige Verständnis des Qualitätsbegriffs*“⁴ hin, denn in der Alltagssprache wird Qualität meist wertend, als Synonym für Güte verwendet (z.B. „Qualitätswein“), obwohl es ursprünglich im Begriff selbst so nicht angelegt ist.

In der **normativen Qualitätsforschung** nimmt man Abstand von einer wertenden Begriffsbestimmung⁵ und bevorzugt eine offene und allgemeine Definition, die auf einer weiteren Übersetzung des lateinischen „qualitas“ beruht: die Beschaffenheit. Demnach ist Qualität die „*Gesamtheit von Merkmalen (und Merkmalswerten) einer Einheit bezüglich ihrer Eignung, festgelegte und vorausgesetzte Erfordernisse zu erfüllen*.“⁶ Dabei interessiert sowohl die Qualität umfangreicher Systeme, als auch einzelne Merkmale oder auch Gruppen davon. Es handelt sich bei Qualität also – noch kürzer formuliert – um die Relation zwischen der realisierten Beschaffenheit einer betrachteten Einheit (oder einzelner Merkmale davon) und der Qualitätsanforderung an diese:

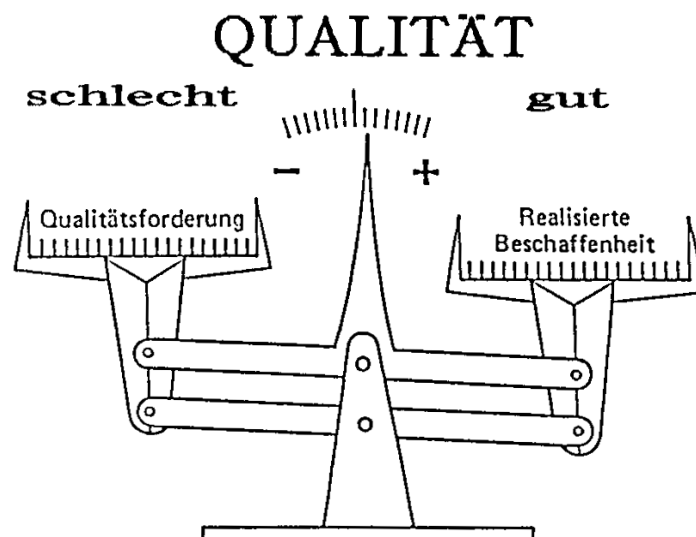


Abb. 1: Der Qualitätsbegriff nach Geiger, S. 46

3 Zur aktuellen Diskussion über den Schulbuchtest der Stiftung Warentest (10/2007) vgl.: Geschichte in Wissenschaft und Unterricht.

4 Geiger, Walter: Qualität als Fachbegriff des QM. In: Zollondz 2001, S. 809.

5 Zollondz 2002, S. 143 ff.

6 ISO 8402 Norm, zit. nach: Becker, S. 9.

Dieser sehr allgemeine und tautologische Sprachgebrauch (Qualität ist, was als Qualität definiert wird) ist methodisch jedoch gefährlich, weil beliebig interpretierbar. Deswegen muss trotz aller Vorbehalte eine konkrete Wertordnung hinter den Forderungen nach Qualität stehen. Kriterien für abschließende Werturteile sind notwendig, um Willkürlichkeit in der Qualitätsbestimmung zu vermeiden (Qualität ist, was der Monopolist definiert).⁷

Im **partialanalytischen Qualitätsverständnis** nach Garvin⁸ können beim Qualitätsbegriff fünf verschiedene Sichtweisen unterschieden werden:

1. Die transzendente Sichtweise – Absolute Qualität

Qualität ist absolut und universell erkennbar, nicht präzise zu definieren und wird nur durch subjektive Erfahrung empfunden.

2. Die produktbezogene Sichtweise – Produktqualität

Qualität ist präzise und messbar, sie ergibt sich aus der Erfüllung allgemein festgelegter Anforderungen, Qualitätsunterschiede werden durch bestimmte Eigenschaften oder Bestandteile eines Produktes auch quantitativ widerspiegelt.

3. Die anwenderbezogene Sichtweise – Kundenqualität

Auf Verwendungszweck und Kundenpräferenz ausgerichtete Betrachtung. Verschiedene Nutzer haben unterschiedliche Wünsche und Bedürfnisse. Die Produkte, die diese Bedürfnisse am besten befriedigen, werden als qualitativ hochwertig angesehen. Kundensubjektiver Bewertungszusammenhang. Beispiel: Ein häufig ausgeliehenes Buch einer Bibliothek.

4. Die prozessbezogene Sichtweise – Prozessqualität

Qualität entsteht durch die richtige Erstellung des Produkts, durch eine gut ausgeführte Arbeit, deren Ergebnis die Anforderungen zuverlässig und sicher erfüllt. Herstellungsorientierter Ansatz. Beispiel: Schweizer Uhrwerk.

5. Die Kosten/Nutzen bezogene Sichtweise – Wertqualität

Qualität wird durch Kosten und Nutzen ausgedrückt. Ein Qualitätsprodukt ist ein Erzeugnis, das einen bestimmten Nutzen zu einem akzeptablen Preis

7 Vgl. Zollondz 2001, S. 813.

8 Garvin, S. 25 ff.

mit einem angemessenen Aufwand erbringt. Wertorientierter Ansatz. Beispiel: IKEA.

Diese Herauslösung einzelner Teilaspekte von Qualität ist bis heute von grundlegender Bedeutung und dient oft besonders im Dienstleistungsbereich zur dankbaren Illustration des Qualitätsbegriffes. Die produkt-, hersteller- und kundenorientierte Sichtweise ist zentraler Bestandteil im Qualitätsmanagement.⁹ Zudem wird in modernen Modellen zum Qualitätsmanagement inzwischen auch die Qualität der Mitarbeiter als sechste Teilqualität ergänzt, unter Berücksichtigung der Arbeitssituation allgemein (Arbeits- und Gesundheitsschutz, Einflussmöglichkeiten am Arbeitsplatz etc.).¹⁰

Im **Bildungs- und Sozialbereich** wird v.a. auf das Begriffsverständnis von Harvey/Green¹¹ zurückgegriffen, das Qualität in fünf Gruppen einteilt. Diese lassen sich einerseits klar unterscheiden, stehen andererseits aber auch miteinander in Beziehung:

1. Qualität als Ausnahme

Qualität wird verstanden als etwas Herausragendes und Besonderes und lässt sich in weiteren drei Varianten unterscheiden:

a) Die traditionelle Vorstellung von Qualität

Qualität wird mit der Vorstellung von etwas sehr Speziellem und Hochklassigem verbunden, das Qualitätsprodukt ist exklusiv und ist für den Besitzer bzw. Verwender ein Statussymbol. Die Qualität kann nicht definiert und gemessen werden, weil dieses Verständnis apodiktisch darauf beruht, dass die Herausgehobenheit selbst schon Qualität ist.

b) Qualität als Exzellenz

Dieses Verständnis erkennt Qualität im Übertreffen von hohen Standards. Es ähnelt dem o.g. traditionellen Konzept, vermeidet jedoch dessen apodiktischen Charakter. Bestandteile von Exzellenz werden zwar benannt, gleichzeitig wird aber sichergestellt, dass sie fast unerreichbar sind. Das Erreichen der Standards ist nur unter sehr spezifischen und

9 Vgl. Zollondz 2002, S. 146.

10 Vgl. ebda.

11 Harvey/Green, S. 17ff.

fast schon unwahrscheinlichen Bedingungen möglich. Die Grundannahme lautet: Die Qualität der Ergebnisse ist eine Funktion der Qualität der Inputs.

c) Qualität als Übereinstimmung mit Standards

Qualität wird einem Produkt zugesprochen, das den vom Erzeuger aufgestellten Minimalstandards entspricht. Es hat eine Reihe von Qualitätskontrollen durchlaufen, die Qualität steigt, wenn die Standards erhöht werden. Als Denkmodell vor allem im Bildungsbereich oft üblich: Qualitätssteigerung bedeutet eine Verbesserung der Gestaltung und der Inhalte von Lehrangeboten sowie ihre Überprüfung.

2. Qualität als Perfektion oder Konsistenz

Der zweite Ansatz konzentriert sich mehr auf die Prozesse, die beim Streben nach Qualität notwendig sind. Auch hier werden zwei – allerdings zusammenhängende – Formen unterschieden:

a) Qualität als Fehlerlosigkeit bzw. Perfektion

Von Anfang an sollen beim Prozess die richtigen Dinge in richtiger Weise getan werden. Perfektion stellt sicher, dass Alles zufriedenstellend verläuft und keine Fehler passieren (Verlässlichkeit). Das Produkt oder die Dienstleistung wird danach beurteilt, ob sie konform ist mit vordefinierten und messbaren Spezifikationen im Sinne von Anforderungen. Diese Spezifikationen sind nicht selbst ein Standard und werden nicht an anderen, externen Standards gemessen.

b) Kultur der Qualität

Alle Mitglieder einer Organisation, alle beteiligten Personen und Instanzen sind für Qualität verantwortlich, nicht nur die Qualitätskontrollen. Damit wird die Verantwortung für Qualität demokratisiert. Eine Organisation ist in einzelne Einheiten aufgegliedert, jede ist ein Knotenpunkt für Qualität, jede Einheit muss ihre Anforderungen spezifizieren und ihr „Produkt“ im gesamten Prozess überprüfen.

3. Qualität als Zweckmäßigkeit

Dieser Ansatz bezieht Qualität auf den Zweck eines Produktes oder einer Dienstleistung. Qualität wird also beurteilt nach dem Ausmaß, in dem ein Produkt seinen Zweck erfüllt. Dabei kann jedes Produkt oder jede Dienst-

leistung potentiell seinen Zweck erfüllen, Perfektion ist dabei irrelevant. Die Frage nach der Zweckmäßigkeit kann zudem aus zwei Perspektiven betrachtet werden:

a) Zweck für den Nutzer

Qualität heißt, die Anforderungen des Konsumenten werden erhoben und erfüllt. Dieses Verständnis ist entwicklungsorientiert, da Zwecke sich ändern können. Im Bildungsbereich ist es nützlich, um die Interessengebundenheit von Qualitätskonzepten deutlich zu machen und auf unterschiedlichen Ebenen zu analysieren.

b) Zweck für die Institution

Qualität bedeutet, eigene festgesetzte Ziele zu erreichen, die Erfüllung des Auftrages der Institution und die Markttauglichkeit der Produkte. Natürlich kann auch hier der Konsument nicht aus dem Auge gelassen werden: Kundenzufriedenheit ist ein klarer Ausdruck von Qualität.

4. Qualität als adäquater Gegenwert

Qualität wird gemessen am Preis-Leistungs-Verhältnis bzw. der Kosten-Nutzen-Relation. Dafür sind einerseits Leistungsindikatoren geschaffen worden: sie sollen Effizienz kenntlich machen (Personal-Studenten-Relationen, Marktanteile usw.). Andererseits wird in so genannten „Consumer Charters“ spezifiziert, was Kunden für ihr Geld erwarten dürfen. Sie sind im öffentlichen und privaten Versorgungsbereich entwickelt worden, um das Fehlen von Marktmechanismen zu kompensieren (Elternrechte, Patientenrechte usw.).

5. Qualität als Transformation

Das transformative Verständnis von Qualität basiert auf der Vorstellung eines qualitativen Wandels, einer grundlegenden Veränderung der Form. Dabei werden zwei Varianten unterschieden:

a) Weiterentwicklung des Teilnehmers

Hier kommt es zu einer Steigerung gegenüber der Ausgangssituation, z.B. in Form eines beruflichen Aufstieges nach einer Fortbildung. Diese Wertsteigerung ist ein Maß für Qualität, das sich auf das Ausmaß bezieht, in dem die Teilnahme an Bildungsmaßnahmen das Wissen, die Fähigkeiten und Fertigkeiten steigert.

b) Ermächtigung der Teilnehmer

Den Teilnehmern wird die Macht gegeben, ihre eigenen Transformationen zu beeinflussen. Diese in der Wirtschaft übliche Strategie dient dazu, Vorteile aus dem Wissen und den Fähigkeiten der Teilnehmer zu ziehen. Im Bildungsbereich wird der Lernende zum Gestalter seines Lernprozesses und übernimmt Verantwortung für die Bestimmung des Stils und der Art von Lernen. Dabei können vier Formen unterschieden werden:

- Evaluation (Erfassung der Lernzufriedenheit)
- Garantie von Minimalstandards und deren Überwachung
- Kontrolle und Organisation des eigenen Lernens (Auswahl von Unterricht und Lernverträgen)
- Entwicklung kritischer Fähigkeiten (Ermächtigung von Bildung nicht nur in der Lernerrolle, sondern auch „für das Leben“)

Zusammenfassung und Fazit: Die hier nur in Auswahl und kurz skizzierten Qualitätsmodelle zeigen die Problematik, ein derartig mehrdimensionales Konstrukt für die eigene Argumentation nutzbar zu machen. Denn je nach Perspektive und Fokus ergeben sich verschiedene Interpretationen und Konzeptionen für das komplexe Phänomen Qualität: Etliche beteiligte Akteure richten ihren jeweils eigenen Blick auf Qualität und was sie darunter verstehen, verschiedene Wissenschaftsbereiche haben unterschiedliche Qualitätsverständnisse und Schwerpunkte, mehrere Qualitätsebenen können untersucht werden (Voraussetzungen, Prozess, Ergebnis, Nutzen):

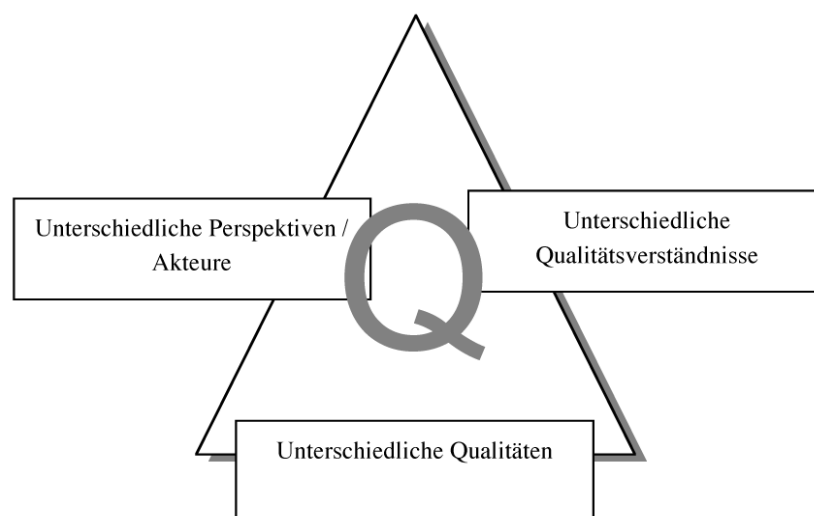


Abb. 2: Dimensionen von Qualität nach Ehlers 2002, S. 4

Das hier zu entwickelnde Qualitätsverständnis, das nicht nur im Kontext eines allgemein pädagogischen Lernbegriffes zu betrachten ist, sondern auch die speziellen Erfordernisse von E-Learning mit einbeziehen soll, wird nicht nur durch einen einzigen der oben ausgeführten Ansätze in Gänze beschrieben. Es stellt vielmehr eine Synthese einzelner Parameter dar. Trotzdem soll der Fokus auf bestimmte Teilaspekte gerichtet werden, die von besonderer Wichtigkeit für die nachfolgenden Ausführungen sind und die einen neuen Blick auf die pädagogische Konstellation „Lerner und Lernangebot“ und deren Gewichtung eröffnen soll: Qualität im Sinne der Anforderungen des Nutzers bzw. Lernalters (anwenderbezogene Sichtweise bzw. Kundenqualität bei Garvin) und Qualität als Transformation durch Ermächtigung des Nutzers bzw. Lernalters (Harvey/Green).

2 Qualität im Bildungsbereich – Paradigmenwechsel, Subjektorientierung und Dienstleistungsverhältnis

Ehlers weist darauf hin, dass v.a. im Sektor der Weiterbildung nicht nur eine Neuorientierung, sondern sogar ein Paradigmenwechsel zu beobachten ist: „[...] *weg von einer Belehrungspädagogik hin zu einer Befähigungspädagogik - von behavioristisch geprägten didaktischen Ansätzen hin zu kognitivistischen und von instruktionalistischen zu konstruktivistischen Ansätzen.*“¹² Es kommt zu einer Individualisierung des Lernprozesses, der Lerner wird zum Nutzer, der seine Bedürfnisse kennt und Lernangebote aktiv und gezielt auswählt. Das lernende Subjekt tritt dabei zunehmend in den Mittelpunkt und erhält die Definitionsmacht für Qualität: Dem Lernenden - seiner Situation, seinen Fähigkeiten, Vorwissen, Zielen und Präferenzen - kommt für gelingendes Lernen konstitutive Bedeutung zu. Lernerbezogene Qualitätsfaktoren stehen nun also im Zentrum der Analyse und Beurteilung der Qualität von Bildungsangeboten.

Um den Begriff des lernenden Subjektes in der Qualitätsdiskussion fest zu etablieren und um die weiteren Überlegungen auch im Bereich des E-Learning theoretisch zu verankern, werden im Folgenden zwei subjektbezogene Forschungsansätze beschrieben, die sich damit befassen, Lernprozesse bzw. Mediennutzungsprozesse aus der Sicht des Lernalters bzw. Nutzers zu betrachten:

- a) Die subjektive Lerntheorie von Holzkamp
- b) Der Nutzen- und Belohnungsansatz

12 Ehlers 2004b, S. 32.

zu a) Die subjektive Lerntheorie von Holzkamp

Holzkamp gilt als Begründer der kritischen Psychologie. Ausgangspunkt der Theorieentwicklung ist die Kritik an traditionellen Lerntheorien, die nach Holzkamp Lernen als ein Lernen unter fremder Kontrolle verstehen und es mit einem reglementierten Lernen gleichsetzen. Eine Analyse der traditionellen Lerntheorien lässt ihn zu dem Schluss kommen, dass Lernen als Problem vom wissenschaftlichen Standpunkt des Lernsubjekts aus nicht vorkommt:

„Dabei ist zu berücksichtigen, dass das Lernsubjekt in den traditionellen Theorien ja nicht real verschwunden sein kann, sondern offenbar nur in irgendeiner Weise theoretisch wegreduziert, verleugnet, damit konzeptionell verfehlt wird.“¹³

Deswegen entwickelt er eine Theorie vom Standpunkt des Subjektes aus, denn aus Sicht der Subjektwissenschaft lebt der Mensch nicht nur unter Bedingungen, sondern ist aktiver Produzent seiner Lebensbedingungen. Mit Beckers lässt sich – stark verkürzt - ein Lernprozess im Holzkampschen Sinne exemplarisch wie folgt beschreiben:

„Ein Subjekt stößt bei einem Handlungsvollzug auf ein Problem. Es erkennt, dass dieses Problem sich durch Lernen lösen lassen kann. Die Handlungsproblematik wird damit vom Subjekt als Lernproblematik identifiziert. Die Lernhandlung kann beginnen, wobei der Handlungsvollzug vorübergehend ausgesetzt wird. Die ursprüngliche Handlung stellt für die Dauer des Lernens eine Bezugshandlung für die ausgegliederte Lernhandlung dar.“¹⁴

Holzkamp spricht von intentionalem Lernen und definiert Lernerfolg als das Erreichen des selbst oder von anderen gesetzten Lernziels. Lernen ist ein Prozess, der im individuellen Lebensinteresse begründet ist und aus individuellen Motiven entsteht.

zu b) Der Nutzen- und Belohnungsansatz

Im Nutzen- und Belohnungsansatz (oder: Uses und Gratifications Approach, oder: Theorie der selektiven Zuwendung) wird der Nutzen von Medien mit ih-

13 Holzkamp, S. 15.

14 Beckers, S. 150.

rer Wirkung in Verbindung gesetzt.¹⁵ Es wird untersucht, warum bestimmte Personen bestimmte Medien benutzen und wie sie von dieser Tätigkeit profitieren. Diese Modell der Mediennutzungsforschung (Was machen Menschen mit den Medien?) hat sich als Gegenströmung bzw. Ergänzung zum Modell der Medienwirkungsforschung (Was machen Medien mit den Menschen?) seit den 1960er Jahren entwickelt, indem es die Rolle des Rezipienten beim Umgang mit Massenmedien in den Blick nimmt. Demnach ist der Rezipient aktiv (also nicht mehr nur passiver Empfänger von Medienbotschaften, sondern eigeninitiatives Subjekt), selektiv (weil er sich nur den Medienbotschaften zuwendet, von denen er sich Befriedigung seiner Bedürfnisse verspricht) und reflektiv (also fähig, seine eigenen Bedürfnisse und Ziele zu erkennen und zu artikulieren). Die Bedeutung eines Medieninhaltes ist folglich kein vorgefertigter Reiz, sondern hängt von der Interpretation des handelnden Individuums ab.

Ehlers spricht beim Thema Qualität im Bildungs- und Sozialbereich von einer Ko-Produktion des Bildungsproduktes: Im Bereich von Lernprozessen treten zwei Akteure gegenüber, die gemeinsam aushandeln müssen, was Qualität ist und die diese Qualität in Interaktionen umsetzen müssen: Der Anbieter und der Kunde (etwa in Gestalt von Lehrenden und Lernenden). Ein ökonomisch verkürzter Kundenbegriff im Sinne von „Qualität ist, was der Kunde will“ ist hier jedoch nicht gemeint. Trotzdem sind Lernprozesse nicht ausschließlich zu erbringende Leistungen eines Bildungsanbieters für einen Kunden, sondern bedürfen der Mitwirkung des Kunden. Auch die Produkte eines Bildungsprozesses sind nicht ausschließlich Ergebnis der Produktionsform oder des Produktionsvorganges der Bildungseinrichtung. Denn was Lernende lernen und können ist in hohem Maße von ihnen selber, ihrem Engagement und ihrer Umwelt bzw. von Sozialisationsprozessen beeinflusst.

Die Vorstellung von einem „Dienstleistungsverhältnis“ und von „Dienstleistungsqualität“ kann nun auch auf den Bereich des Lehrens und Lernens übertragen werden. Danach muss der „Kunde“ in gewissem Maße bei der Erstellung der angeforderten Leistung mitwirken, ja er hat sogar große Mitverantwortung beim Gelingen der Dienstleistung. Darüber hinaus gilt es, die institutionellen, lehrorganisatorischen und staatlichen Rahmenbedingungen zu berücksichtigen.

15 Vgl. Hugger, S. 173 ff.

Für den Bildungsbereich hat dieses Verständnis von Qualität und vom Verhältnis von Anbieter und Kunde entscheidende Konsequenzen¹⁶:

- Betrachtet man ein Bildungsangebot als Dienstleistungsangebot kann Qualität nicht bereits im Voraus gesichert werden und ist erst nach Inanspruchnahme sichtbar und überprüfbar.
- Qualität drückt sich durch eine Zustandsänderung (Transformation) aus und nicht durch eine Eigentumsübertragung.
- Der Empfänger der Dienstleistung ist bei Bildungsangeboten aktiv an der Erstellung beteiligt.

Qualitätsbeurteilungen von E-Learning-Angeboten erfordern demnach eine neue Grundkategorie, die Lernerorientierung. Dieser Perspektivenwechsel ist auch mit Konsequenzen für die empirische Forschung verbunden.

3 Lernerorientierung als Imperativ – Konsequenzen für die empirische Forschung

Evaluation ist ein Verfahren, mit dem Qualität gemessen werden kann und dient zunehmend häufiger auch im Bildungsbereich der Qualitätssicherung von etablierten pädagogischen Systemen und Programmen.¹⁷ Tergan umfasst in seiner Definition von Evaluation v.a. Bildungsangebote und deren Beurteilung:

„Evaluation ist die systematische und zielgerichtete Sammlung, Analyse und Bewertung von Daten zur Qualitätssicherung und Qualitätskontrolle. Sie gilt der Beurteilung von Planung, Entwicklung, Gestaltung und Einsatz von Bildungsangeboten bzw. einzelner Maßnahmen dieser Angebote (Methoden, Medien, Programme, Programmteile) unter den Aspekten von Qualität, Funktionalität, Wirkungen, Effizienz und Nutzen.“¹⁸

Der empirische Forschungsstand zu lernerbezogenen Qualitätsanforderungen im Bereich des E-Learning innerhalb der Musikpädagogik ist aber noch recht dürftig. Bei den zahlreichen Kriterienkatalogen zur Überprüfung der

16 Vgl. Ehlers 2004a, S. 56 f.

17 Vgl. Krapp/Weidenmann, S. 29.

18 Tergan, S. 23.

Lernwirksamkeit von Software wird das aktiv beteiligte Lernsubjekt kaum berücksichtigt.¹⁹

Zwei Forderungen sollen als Konklusion aus den vorangegangenen Ausführungen neue Impulse für das empirische Forschen im Bereich des E-Learning geben:

- a) Von der Angebotsorientierung zur Nutzerorientierung!
- b) Von der summativen zur formativen Evaluation!

zu a) Von der Angebotsorientierung zur Nutzerorientierung!

Evaluation darf sich nicht ausschließlich auf die Aufbereitung des Inhaltes beziehen. Die pädagogisch-didaktische Konzeption ist zwar wichtige Voraussetzung für erfolgreiches Lernen, doch der Nutzer und sein Umgang mit dem Angebot, das lernende Subjekt, ist ebenso wichtiger didaktischer Ratgeber bei der Entwicklung. Die Frage ist nicht mehr „Welche Inhalte können wie am besten programmiert werden“, sondern „Welche Methode bringt für den individuellen Nutzer den größtmöglichen Lernerfolg?“ Der Primat der Technologieorientierung muss dem Primat der Technologieunterstützung weichen. Denn E-Learning-Angebote sind, wie alle Unterrichts- bzw. Lernmethoden, nicht für alle Inhalte, Ziele und Lernertypen sinnvoll: *„Nicht alle Lerner profitieren gleichermaßen vom computerunterstützten Lernen und nicht alle Inhalte eignen sich in gleichem Maße dazu, über [...] Lernprogramme vermittelt zu werden.“*²⁰ Qualität entsteht beim Lernen, bei der Interaktion von Lerner und Lernarrangement, sie ist ein gemeinsames Bedürfnis und eine gemeinsame Produktion von „Anbieter“ und „Nutzer“. Deswegen müssen bei der Evaluierung von E-Learning-Angeboten auch beide „Partner“ in gleicher Weise und Intensität in den Mittelpunkt der Untersuchung rücken.

zu b) Von der summativen zur formativen Evaluation!

Es werden grundsätzlich zwei Typen von Evaluation unterschieden, die formative und die summative Evaluation: *„Formative Evaluation dient der Qualitätssicherung. Ihr vorrangiges Ziel ist die Ermittlung von Schwachstellen. Sie erfolgt zumeist entwicklungsbegleitend und liefert Daten, die für die Optimie-*

19 Vgl. z.B. www.sodis.de oder www.i-cd-rom.de (18.02.2010), vgl. auch: Holzinger.

20 Reinmann-Rothmeier/Mandl, S. 635.

*rung der Gestaltung eines Bildungsangebotes verwendet werden.*²¹ Diese Form der Qualitätssicherung ist also prozessorientiert. Sie wird z.B. bei der Überprüfung von Lernprogrammen im Entwicklungsstadium, also bei Prototypen und Vorversionen eingesetzt. Probanden werden bei der Programmbearbeitung beobachtet und befragt, die Ergebnisse fließen in die weitere Entwicklung und in die Verbesserung des Programms ein. Es wird überprüft, ob das verwendete System effizient ist, also für den geplanten Zweck richtig gestaltet ist. Empirische Erhebungsmethoden sind Lernerprotokolle, Interviews und Expertenratings.

Von der formativen Evaluation wird die summative Evaluation unterschieden: *„Summative Evaluation dient der Kontrolle von Qualität, Wirkungen und Nutzen eines Bildungsangebots. Das Interesse gilt der Frage, ob ein Bildungsangebot bzw. einzelne Komponenten des Angebotes bestimmten Erwartungen in der praktischen Anwendung gerecht werden.“*²²

Diese Form der Qualitätssicherung ist ergebnisorientiert. Sie wird zur Beschreibung und Bewertung von bereits abgeschlossenen Lernprogrammen, wie sie z.B. im Handel erhältlich sind, eingesetzt. Sie dient zur Legitimation von Programmen und liefert Daten über positive und negative Bedingungen sowie Verbesserungsmöglichkeiten bei deren Einsatz. Die Ergebnisse können allerdings auch zur Weiterentwicklung des Programms oder seines Einsatzes dienen (z.B. in Folgeversionen und Updates). Es wird überprüft, ob das verwendete System effektiv ist, also die richtige Lösung für die gegebenen pädagogische Aufgabenstellung darstellt. Methoden der summativen Evaluation sind vor allem Wirkungsanalysen. Sie untersuchen die Akzeptanz des Programms durch den Lerner, ablaufende Lernprozesse und den Lernerfolg, sowie die Fähigkeiten zum Lerntransfer.

Bei einer stärkeren Einbeziehung des lernenden Subjektes reicht es also nicht aus, *Lernergebnisse* mit Hilfe von Tests und Punktesystemen zu ermitteln, sondern den aktiven Nutzer und seinen *Lernprozess* mit einzubeziehen. Bereits bei der Entwicklung von E-Learning-Angeboten muss der Lerner selbst zum wichtigen Korrektiv und zur letztlich entscheidenden Instanz bezüglich der Qualitätsbeurteilung des „Produktes“ werden!

21 Tergan, S. 25. (Hervorhebungen im Original)

22 Ebda. (Hervorhebungen im Original)

4 Lerner integrierende Qualitätsbeurteilung am Beispiel von Musik-Lernsoftware

Vom Verfasser liegt eine Untersuchung vor, die zum ersten Mal explizit das lernende Subjekt in die Qualitätsbeurteilung eines E-Learning-Angebotes aus dem musikpädagogischen Bereich mit einbezieht.²³ Grundlage ist ein Evaluationsmodell von Reinmann-Rothmeier/Mandl/Prenzl das sich vor allem zur Evaluierung von Lernsoftware im universitären Bereich bewährt hat.²⁴ Zentrale Elemente dieses Modells sind die der Qualitätsanalyse durch Experten und der Wirkungsanalyse durch den Lerner selbst.

Die **Qualitätsanalyse** ist Teil der formativen Phase im Evaluationsmodell. Experten in den Bereichen Ergonomie, Inhalt und Didaktik beurteilen den Evaluationsgegenstand, die daraus gewonnenen Erkenntnisse fließen in die endgültige Programmversion oder in verbesserte Versionen des Programms mit ein. Diese „Außenansicht“ bei der Qualitätsbeurteilung soll als wichtiger Faktor bei der Evaluation natürlich nach wie vor seine Berechtigung haben. Lediglich die Ausschließlichkeit dieser Vorgehensweise soll zugunsten einer zusätzlichen „Innenansicht“ des eigentlich handelnden Subjektes aufgehoben werden. Dies geschieht im beschriebenen Modell durch die Hinzunahme der sog. Wirkungsanalyse, die sowohl in der formativen wie auch in der summativen Phase einer Evaluation von Bedeutung ist.

Die **Wirkungsanalyse** umfasst folgende Dimensionen:

- Die Dimension **Akzeptanz** erfasst die allgemeine Zufriedenheit des Lernalers mit der Lernumgebung. Nur wenn der Lernende die Lernumgebung akzeptiert und sich grundsätzlich von ihr Wissenserwerb verspricht, ist auch ein günstiger Lernprozess und ein guter Lernerfolg möglich. Die allgemeine Zufriedenheit ist ebenso eine wesentliche Voraussetzung für die regelmäßige Nutzung eines Lernangebotes. Die grundsätzliche Frage lautet: Wie wird die Lernumgebung von den Lernern aufgenommen? Wie zufrieden sind die Lernenden mit dem Lernangebot?
- In der **Lernprozessanalyse** wird die Aufnahme und Verarbeitung der Lerninhalte untersucht. Vorgang und Qualität des Lernens aus der Lernerperspektive stehen im Mittelpunkt. Die grundsätzliche Frage lautet: Wie wer-

23 Vgl. Stubenvoll.

24 Vgl. Reinmann-Rothmeier/Mandl/Prenzl, S. 267 ff.

den die Lerninhalte von den Lernenden aufgenommen und verarbeitet? Welche Strategien entwickeln sie im Umgang mit dem Lernangebot? Der Analyse von emotionalen, motivationalen und kognitiven Aspekten wird zusätzlich eine besondere Bedeutung zugewiesen. Affektive und motivationale Variablen sind dabei von gleicher Wichtigkeit wie die kognitiven Bedingungen.

- Die **Lernerfolgs-** und **Transferanalyse** untersucht die Ergebnisse eines Lernprozesses. Lernerfolg kann sich in Erwerb bzw. Veränderung von Verständnis, Wissen, Fertigkeiten und Einstellungen zeigen. Zentrale Fragen sind: Was haben die Teilnehmer gelernt? Wurde das Lernziel erreicht? Dabei geht es nicht um die Abfrage von einzelnen inhaltlichen Fakten im Sinne eines objektiven Testverfahrens, sondern um die subjektive Einschätzung des eigenen Lernerfolges durch den Lerner selbst.

Das Element der Wirkungsanalyse ist ein ideales Instrument zur Lerner integrierenden Qualitätsbeurteilung, weil das Modell damit dem aktuellen Paradigmenwechsel vom angebotsorientierten zum subjektorientierten Qualitätsbegriff Rechnung trägt, indem es konkret und direkt den „Endkunden“, nämlich das mit dem Programm lernende Subjekt, befragt.

Das vom Autor veränderte und erweiterte „Münchener Modell“ ist flexibel modifizierbar und einsetzbar. Es liefert differenzierte Bewertungen und Aussagen, die sowohl für den Lerner als auch für den Lehrer bzw. Anbieter von E-Learning-Angeboten pädagogisch nutzbar gemacht werden können.

Literatur

Becker, Peter: *Prozessorientiertes Qualitätsmanagement*. Renningen: expert-Verlag, 2001.

Beckers, Erich: *Grundlegung des Konstrukts „Subjektive Theorie musikalischen Lernens*. In: Pfeffer/Vogt, S. 139-154.

Ehlers, Ulf (2002): *Qualität beim E-Learning: Der Lernende als Grundkategorie bei der Qualitätssicherung*. In: *Medienpädagogik*. Zeitschrift für Theorie und Praxis der Medienbildung. Onlinepublikation vom 25.03.2002. www.medienpaed.com/02-1/ehlers1.pdf. (23.08.2009)

Ehlers, Ulf-Daniel (2004a): *Qualität im E-Learning aus Lernaltersicht*. Grundlagen, Empirie und Modellkonzeption subjektiver Qualität. Wiesbaden: VS Verlag für Sozialwissenschaften, 2004.

- Ehlers, Ulf-Daniel (2004b): *Erfolgsfaktoren für E-Learning: Die Sicht der Lernenden und mediendidaktische Konsequenzen*. In: Tergan/Schenkel, S. 29-49.
- Friedrich, H. F./ Eigler, G./Mandl, H./Schnotz, W./Schott, F./Seel, N.: *Multi-mediale Lernumgebungen in der betrieblichen Weiterbildung*. Gestaltung, Lernstrategien und Qualitätssicherung. Neuwied: Luchterhand, 1997.
- Garvin, David: What does Product Quality Really Mean? In: Sloan Management Review. Vol. 25, Fall 1984. S. 25-43.
- Geiger, Walter: *Qualitätslehre*. Einführung, Systematik, Terminologie. Braunschweig/Wiesbaden: Vieweg, ²1994.
- Geschichte in Wissenschaft und Unterricht. Schulbücher – Analyse und Kritik*. Nr. 60, 01/2009. Seelze: Erhard Friedrich Verlag. (Zeitschrift)
- Harvey, Lee/Green Diana: *Qualität definieren*. Fünf unterschiedliche Ansätze. In: Helmke/Hornstein/Terhart, S. 17-39.
- Helmke, Andreas/Hornstein, Walter/Terhart, Ewald (Hrsg.): *Qualität und Qualitätssicherung im Bildungsbereich: Schule, Sozialpädagogik, Hochschule*. Weinheim: Beltz, 2000. (Zeitschrift für Pädagogik, 41. Beiheft)
- Holzinger, Andreas: *Basiswissen Multimedia*. Würzburg: Vogel, 2001. Bd. 3: Design.
- Holzkamp, Klaus: *Lernen*. Subjektwissenschaftliche Grundlegung. Frankfurt: Campus, 1993.
- Hugger, Kai-Uwe: *Uses-und-Gratification-Approach und Nutzenansatz*. In: Sander/Gross/Hugger, S. 173-177.
- Krapp, Andreas/Weidenmann, Bernd: *Pädagogische Psychologie*. Ein Lehrbuch. Weinheim: Beltz, ⁴2001.
- Pfeffer, Martin/Vogt, Jürgen (Hrsg.): *Lernen und Lehren als Thema der Musikpädagogik*. Sitzungsbericht 2002 der Wissenschaftlichen Sozietät Musikpädagogik. Münster: Lit Verlag, 2004.
- Reinmann-Rothmeier, G./Mandl, H./Prenzel, M.: *Qualitätssicherung bei multimedialen Lernumgebungen*. In: Friedrich/Eigler/Mandl/Schnotz/Schott/Seel, S. 267 - 333.
- Reinmann-Rothmeier, Gabi/Mandl, Heinz: *Unterrichten und Lernumgebungen gestalten*. In: Krapp/Weidenmann, S. 601-646.

- Sander, Uwe/von Gross, Friederike/Hugger, Kai-Uwe (Hrsg.): *Handbuch Medienpädagogik*. Wiesbaden: VS Verlag für Sozialwissenschaften, 2008.
- Schenkel, Peter/Tergan, Sigmar-Olaf/Lottmann, Alfred (Hrsg.): *Qualitätsbeurteilung multimedialer Lern- und Informationssysteme*. Evaluationsmethoden auf dem Prüfstand. Nürnberg: BW-Verlag, 2000.
- Stubenvoll, Matthias: *Musiklernen am Computer*. Zur Qualität von Musik-Lernsoftware und ihrer empirischen Überprüfung. Essen: Die Blaue Eule, 2008.
- Tergan, Sigmar-Olaf: *Grundlagen der Evaluation: Ein Überblick*. In: Schenkel/Tergan/Lottmann, S. 22-51.
- Tergan, Sigmar-Olaf/Schenkel, Peter (Hrsg.): *Was macht E-Learning so erfolgreich?* Grundlagen und Instrumente der Qualitätsbeurteilung. Berlin: Springer, 2004.
- Zollondz, Hans-Dieter (Hrsg.): *Lexikon Qualitätsmanagement*. Handbuch des modernen Managements auf der Basis des Qualitätsmanagements. München: Oldenbourg, 2001.
- Zollondz, Hans-Dieter: *Grundlagen Qualitätsmanagement*. Einführung in Geschichte, Begriffe, Systeme und Konzepte. München: Oldenbourg, 2002.

Darf es noch ein bisschen mehr sein?

Auswirkungen von Stimmtraining bei Chorklassenkindern

1 Einleitung/Forschungsstand

Chorsingen als gesellschaftlich weit verbreiteter Ausdruck von Musikkultur ist ein vielfältig untersuchter Forschungsgegenstand. Aus unterschiedlichen Forschungsdisziplinen existieren zahlreiche wissenschaftliche Studien zur Wirkung und Bedeutung des Chorsingens sowohl in Laien- als auch in Spitzenchören (Überblick siehe: Kreutz, 2004). Beispielsweise Ternström (1991) führte psychophysikalische Untersuchungen an Chören durch, um Mechanismen der Stimmproduktion mit Aspekten der Raumakustik und psychoakustischen Vorgängen innerhalb der auditiven Verarbeitung zu verbinden. Des Weiteren wurden in einigen Studien sozial- und gesundheitspsychologische Auswirkungen des Chorsingens eruiert (z.B. Clift & Hancox 2001; Bailey & Davidson, 2003). Hinsichtlich psychophysiologischer Auswirkungen geben einige Forschungsstudien Hinweise auf positive Einflüsse gesanglicher Aktivitäten von Laien auf das Immunsystem (z.B. Kreutz, Bongard, Grewe u.a., 2004).

Bedarfe zur Förderung des Singens bei Kindern werden seit einigen Jahren nicht nur seitens musikpädagogischer Fachliteratur erkannt und diskutiert (z.B. Hoos de Jokisch, 2003; von Gutzeit, 2001), sondern führten unter anderem zu bildungspolitischen Initiativen, wie etwa der Einrichtung so genannter “Chorklassen” nach dem Vorbild von “Bläser-“ und “Streicherklassen” in Niedersachsen. Chorklassenkinder erhalten im Rahmen des zweistündigen Regelunterrichts Musikstunden, in denen die Unterrichtsinhalte unter Akzentuierung des Chorsingens erarbeitet werden. Diese Musikstunden werden durch eine wöchentlich stattfindende halbe Stunde Stimmbildung in Kleingruppen er-

1 geb. Freudenhammer

gänzt. Ziel ist eine vokalpädagogische Breitenbildung im Primarbereich und in der Sekundarstufe I.

Ein in diesem Zusammenhang wenig beachtetes Phänomen betrifft hohe Prävalenzraten von Stimmstörungen im Kindes- und Jugendalter, deren vermuteter Anteil in der medizinischen Forschung zwischen 6 und 24 Prozent angegeben wird (Connor, Cohen, Theis u.a., 2008). Diese Beobachtungen sind angesichts der allgemeinen Bedeutung der Funktionalität des Stimmapparats für die kindliche Entwicklung alarmierend. Sallat (2009) erläutert, dass Störungen des Stimmapparats vielfältige Lernprozesse beeinträchtigen können. Betroffene Kinder „gelten als Risikokinder ersten Ranges für die Ausbildung von Folgeproblemen im schulischen, kognitiven und sozial-emotionalen Bereich.“ (Sallat 2009, S.85)

Untersuchungen zeigen, dass Kinder mit pathologischem Stimmklang anhand von Aufnahmen der Stimmen sowohl von Erwachsenen als auch von Gleichaltrigen eher mit negativ besetzten Assoziationen und Stereotypen bedacht werden im Vergleich zu Kindern mit gesunden Stimmen (Ruscello, Lass & Podbesek, 1988; Lass, Ruscello, Stout, & Hoffmann, 1991). Leitfadeninterviews mit Kindern und Jugendlichen verschiedener Altersgruppen zwischen 2 und 18 Jahren bzw. ihren Familien legen nahe, dass Betroffene von Stimmstörungen in Folge eines pathologischen Stimmklangs körperlichem, sozialem und emotionalem Leidensdruck ausgesetzt sein können (Connor, Cohen, Theis u.a., 2008). Somit erhebt sich die Frage inwiefern gezielte gesangliche Förderung, wie es etwa in Konzepten von Chorklassen vorgesehen ist, der Prävalenz von Stimmstörungen und ihren Folgen entgegen wirken könnte.

Es ist zu vermuten, dass musikalische Stimmförderung für die Stimmentwicklung im Allgemeinen und den Spracherwerb im Besonderen positive Potentiale in sich bergen kann. Neurowissenschaftliche Untersuchungen zeigen, dass das menschliche Gehirn sprachliche und musikalische Syntax teils in identischen Hirnregionen verarbeitet (Jentschke, Koelsch, Sallat u.a., 2008). Daher scheinen Wechselwirkungen zwischen Musik- und Sprachlernen im Entwicklungsverlauf wahrscheinlich. In der Tat zeigten Studien von Sallat (2009) mögliche Bedeutungen von musikalischem Training als Intervention in der Sprachtherapie. Dieser Autor meint, dass sprachgestörte Kinder vom Erwerb musikalischer Expertise hinsichtlich des Spracherwerbs nachhaltig profitieren. Gleichwohl bleibt bislang offen, inwiefern gerade Gesangsexpertise die Entwicklung vokaler Fähigkeiten und Fertigkeiten fördern kann.

Frühere Studien zeigen, dass eine gezielte Förderung der Singstimme einen positiven Einfluss auf die Stimmentwicklung und die Beherrschung der

Stimmwerkzeuge haben kann. Vokalen Leistungen von sängerisch aktiven im Vergleich zu sängerisch nicht aktiven Menschen wurde in empirischen Studien speziell hinsichtlich der Stimmleistungsparameter Tonhöhe und Lautstärke nachgegangen. Querschnitt-Vergleiche zwischen Sängern und Nicht-Sängern offenbarten in der ersten Gruppe vergleichsweise größere Tonhöhen- und Dynamikumsfänge und damit eine größere Leistungsfähigkeit der Stimme bei Erwachsenen (z.B. Awan, 1991; Sulter, Schutte & Miller, 1995; Åkerlund & Gramming, 1994) sowie auch bei Kindern (Berger & Walde, 2002; Fuchs, Heide, Hentschel u.a., 2006). Auch gibt es Hinweise darauf, dass bei Kindern mit spezieller Ausbildung der Singstimme die Aperiodizität der Stimmlippen-schwingung während der Phonation geringer ist als bei Kindern ohne vokalpädagogische Betreuung (Dejonckere, Wieneke, Bloemenkamp & Lebacqz, 1996). Dies indiziert eine präzisere biomechanische Kontrolle über den Stimmapparat.

Längsschnittstudien zur Entwicklung von Stimmparametern liegen nur in geringem Umfang vor. Le Borgne & Weinrich (2002) sowie Mendes, Rothman, Sapienza & Brown (2003) berichten von signifikant erweiterten Tonhöhenumsfängen und verbesserter dynamischer Kontrolle über die Singstimme bei Erwachsenen, die über mehrere Monate an intensivem Stimmtraining teilnahmen. Vergleichbare Interventionsstudien bei Kindern liegen unseres Wissens bislang nicht vor.

Die vorliegende Längsschnittstudie wurde konzipiert, um Hinweise auf mögliche präventive Wirkungen von Stimmtraining auf die vokale Gesundheit bei Kindern der fünften Klasse aufzudecken. Vorab sei jedoch ausdrücklich darauf hingewiesen, dass die Untersuchung nicht geeignet ist, entsprechende Wirkungszusammenhänge nachzuweisen. Im Längsschnitt-Vergleich der Entwicklung von Kindern in Chorklassen mit Kindern, die am Regelunterricht Musik teilnehmen, kann des Weiteren die Effektivität des vokalpädagogischen Förderprogramms "Chorklasse" bezüglich der Stimmleistung und Stimmqualität von Kindern überprüft werden. Vor diesem Hintergrund beinhaltet das vorliegende Forschungsdesign die Erhebung unterschiedlicher Stimmparameter, die in der medizinisch-phoniatrischen Forschung als behaviorale Indikatoren der Stimmleistung, Stimmqualität und Stimmgesundheit verwendet werden.

2 Methoden

Das Design der hier vorliegenden Studie sieht zwei unabhängige Kohorten vor, die jeweils im Längsschnitt untersucht werden. Die Kohorten unterschei-

den sich bezüglich des sozio-demographischen Hintergrundes der teilnehmenden Schülerinnen und Schüler und hinsichtlich der Schulform (integrierte Gesamtschule vs. Gymnasium). In den vorliegenden Analysen wird lediglich das Datenmaterial eines Schuljahres der ersten Kohorte einbezogen, die an einer Gesamtschule in Oldenburg in einem sozio-ökonomisch "schwächeren" Einzugsgebiet erhoben wurde als es in der zweiten Kohorte an einem Gymnasium in Hannover vorliegt (diese Ergebnisse werden in späteren Publikationen behandelt).

2.1 *Probanden*

Innerhalb der ersten Kohorte nahmen insgesamt fünfzig Fünftklässler (27 Mädchen und 23 Jungen) einer integrierten Gesamtschule in Oldenburg innerhalb zweier Probandengruppen teil: Schülerinnen und Schüler einer Chorklasse (Gruppe „Chorklasse“, $n = 32$) und einer Vergleichsgruppe (Gruppe „Regelunterricht“, $n = 18$). Die Einteilung der Gruppen erfolgte durch ein teilweise selbstselektiertes Verfahren (dies wird im Abschnitt "Einschränkungen der Studie" genauer diskutiert). Im ersten Teil der Tabelle 1 wird die Alters- und Geschlechtsverteilung beider Gruppen detailliert aufgeführt.

Alle Kinder beider Probandengruppen erhielten zwei Stunden Musikunterricht (à 45 Minuten) wöchentlich. Im Musikunterricht der Chorklassenkinder wurden die Lerninhalte unter Akzentuierung des Singens - vor allem des Chorsingens - erarbeitet und durch eine wöchentlich stattfindende halbe Stunde Stimmbildung in Kleingruppen (4 bis 6 Kinder) ergänzt, in der vokalpädagogische Arbeit im Vordergrund stand. Dieser Stimmbildungsunterricht in Kleingruppen wurde von Musikstudentinnen und Musikstudenten der Universität Oldenburg erteilt, die in ihrem Lehramtsstudium den Schwerpunkt auf die Bereiche Kinderstimmbildung und Kinderchorleitung gelegt hatten und in der Anfangsphase des Unterrichtens Supervision von einer Person mit langjähriger Erfahrung mit Kinderstimmbildung erhielten. Die Probanden mit Regelunterricht Musik, der ebenfalls Klassensingen, jedoch keinen speziellen Stimmbildungsunterricht beinhaltete, erhielten statt der Stimmbildung in Kleingruppen zusätzlich eine halbe Stunde Zeit zum freien Spielen.

Die Vergleichbarkeit der beiden Probandengruppen wurde anhand der Kriterien Alter, Geschlecht, Schulleistungen, Vorerfahrungen im Chorsingen und Bildungsstatus der Eltern überprüft (Tabelle 1). Chi-Quadrat-Tests bzw. T-Tests ergaben, dass die Verteilung zwischen den Gruppen für diese Variablen vergleichbar ist.

Eine Quantisierung der Schulleistungen erfolgte auf der Grundlage der individuellen Schulempfehlungen für die weiterführende Schule (HS/RS = Hauptschule/Realschule, Gy = Gymnasium), da an dieser integrierten Gesamtschule in der fünften Jahrgangsstufe keine Schulnoten vergeben werden. Über Befragungen der Eltern bzw. der Erziehungsberechtigten wurden musikalische Hintergründe der Probanden eruiert. Von den Eltern berichtete Singaktivitäten innerhalb der Familie waren in beiden Probandengruppen vernachlässigbar und wurden in Analysen zur Vergleichbarkeit der Gruppen nicht einbezogen.

Der sozio-ökonomische Status der Eltern bzw. Erziehungsberechtigten wurde anhand eines im Rahmen von Pisa erprobten Verfahrens (Züchner, Arnoldt, & Vossler, 2008), dem “International Standard Classification of Education” (ISCED), ermittelt. Im Zusammenhang der vorliegenden Studie wurden zwei Stufen unterteilt: Level I (niedrig) umfasste die Eltern, die in die ISCED-Kategorien 0 bis 4 fielen; Level II (hoch) die Eltern, die in die ISCED-Kategorien 5 bis 6 fielen (siehe Tabelle 1 folgende Seite).

Die Befragung der Eltern umfasste ebenso die Auskunft, ob die Probanden im Zeitraum der Untersuchungen Stimmkrankheiten hatten. Lediglich in der Gruppe der Chorklassenkinder berichteten die Eltern eines Mädchens von einer infektiösen Entzündung der Stimmlippen die eine Woche andauerte und ohne medikamentöse Behandlung wieder abklang. Da in der Gruppe mit Regelunterricht die Eltern dreier Kinder berichteten, dass ihre Kinder Asthmatiker sind, wurden diese Probanden von den Analysen ausgeschlossen, worin die Probandenanzahl von $n = 18$ (statt ursprünglich $n = 21$) resultierte. Die Datensätze bezüglich der Schulempfehlungen und der Bildungslevel der Eltern konnten fast vollständig mit lediglich jeweils einem fehlenden Wert pro Kategorie erhoben werden.

	Chorklasse	Regelunterricht
Geschlecht	20 Mädchen, 12 Jungen	7 Mädchen, 11 Jungen
Durchschnittsalter	10;9	10;11
Schulempfehlung		
<i>HS/RS</i>	62,5 %	64,71 %
<i>Gy</i>	37,5 %	35,29 %
Vorerfahrung Chor	46,88 %	33,34 %
Bildungslevel Eltern		
<i>Level I (niedrig)</i>	51,6 %	50 %
<i>Level II (hoch)</i>	48,4 %	50 %

Tabelle 1: Merkmale der Probandengruppen

Hinweis: Das Durchschnittsalter beider Probandengruppen ist in Jahren und Monaten (x;y) angegeben. HS/RS = Hauptschul-/Realschulempfehlung, Gy = Gymnasialempfehlung (genauere Erläuterungen siehe Text).

2.2 Erhebungen

Stimmfeldmessungen und die Bestimmung des Dysphonia Severity Index (DSI) wurden bei allen Probanden zu Beginn und Ende des Schuljahres in Einzelsitzungen durchgeführt. Das gesamte Design der Studie umfasst eine multidimensionale Testbatterie, die weiterhin Inventare zum Selbstkonzept, zur Emotionsregulation, zur Selbstwahrnehmung der Stimme etc. beinhaltet. Die Ergebnisse dieser Erhebungen werden in zukünftigen Publikationen dargestellt.

2.2.1 Stimmfeldmessungen

Die Stimmfeldmessung (auch bekannt als Phonetogramm oder Stimmumfangprofil) ist ein standardisiertes Verfahren zur Messung des Tonhöhen- und Dynamikumfangs von Stimmen. Durch dieses Untersuchungsverfahren können die stimmlichen Leistungsgrenzen hinsichtlich der Lautstärke und Tonhöhe einer Stimme objektiviert werden. Stimmfeldmessungen werden in der Fachliteratur als sinnvolles Werkzeug zur Erfassung von Stimmleistungen sowohl

bei Kindern als auch bei Erwachsenen beschrieben (Gramming, 1988; Schutte & Seidner, 1983; Coleman, 1993; Åkerlund & Gramming, 1994; Awan, 1991; Sulter u.a., 1995; Le Borgne & Weinrich, 2002; Böhme & Stuchlik, 1995; Wuyts, Heyler, Mertens u.a., 2002).

Ein Stimmfeld besteht aus zwei Kurvenzügen, die die Stimmintensität der lautesten und leisesten Stimmgebung als eine Funktion der Grundfrequenz (F_0) über den gesamten Stimmumfang hinweg abbilden. Nach phoniaterischer Konvention wird der Frequenzumfang in einem Diagramm auf der horizontalen Achse dargestellt, während die Stimmintensität auf der vertikalen Achse des Stimmfeldes aufgetragen wird. Dadurch gibt die Stimmfeldmessung nicht nur Auskunft über die stimmlichen Leistungen hinsichtlich der Tonhöhe und Lautstärke, sondern auch über die Interaktion dieser Parameter (Wuyts u.a., 2002).

Stimmfeldmessungen der Singstimme liefern wichtige Informationen über den Tonhöhen- und Dynamikumfang der Stimme und stellen daher ein wichtiges Instrument zur Erfassung der Stimmleistungen dar. Durch das präzise Erfassen der physiologischen Grenzen einer Stimme können Stimmfeldmessungen sinnvollen Einsatz in der Quantifizierung der stimmlichen Leistungsfähigkeit und in der Dokumentation von Ausbildungsverläufen finden.

In der vorliegenden Studie wurden die Stimmfeldmessungen mit dem Mess-System DiVAS Stimmdiagnostik v2.1 (XION medical GmbH, Berlin) durchgeführt. Die Aufnahme-prozedur folgte dabei im Wesentlichen den Empfehlungen der Union der Europäischen Phoniater ("Union of European Phoniatricians") (Schutte & Seidner, 1983).

Alle Probanden wurden vor den Erhebungen über den Ablauf der Mess-prozedur informiert. Datenerhebungen fanden in einem ruhigen und separaten Raum innerhalb des Schulgebäudes statt, der keine speziellen akustischen Bedingungen aufwies ("Wohnzimmerakustik"). Dies ermöglichte den Probanden während der Phonation durch Reflexionen des Raumes eine bessere auditive Selbstkontrolle als etwa in einer Audiometriekabine (Schutte & Seidner, 1983).

Die Probanden nahmen für die Messungen eine locker stehende Haltung ein, die für das Singen als physiologisch günstig erachtet wird (Nawka, Franke & Galkin, 2006). Die Aufnahmen erfolgten durch das "XION-Mikrofon-Headset", welches mit einer 30 cm langen Stange versehen ist. Am Ende der Stange befindet sich das Schallpegelmessgerät, um eine feste Position des Mikrofons zu garantieren und um den Probanden zu ermöglichen, den Kopf

frei zu bewegen. Durch eine integrierte Elektronik ist der Mikrofonanschluss automatisch kalibriert, wodurch reproduzierbare Bedingungen gegeben sind. Die Grundfrequenz wird in Hertz (Hz), der Stimmeschallpegel in Dezibel (dB(A)) erfasst und alle Daten werden in einem Computerprogramm gespeichert. Alle Stimmfeldmessungen wurden von der Erstautorin durchgeführt, die über langjährige Erfahrung mit Kinderstimmbildung verfügt.

Im Verlauf der Messungen sang jeder Proband einmal mit der leisesten und schließlich mit der lautesten Stimmgebung über den gesamten möglichen Tonhöhenumfang vom tiefsten bis zum höchsten möglichen Ton, wobei alle Töne von der Versuchsleiterin vorgesungen wurden. Dabei wurde der Testvokal [a:] verwendet, da dieser aus akustischen Gründen die höchsten und niedrigsten Intensitäten der Stimme messbar macht. Bei den Aufnahmen wurde immer mit der leisestmöglichen Phonation begonnen, da Probanden mit ungeübten Stimmen häufig nach der lautestmöglichen Stimmgebung nicht schnell genug entspannen um wieder sehr leise Töne erzeugen zu können. (Nawka, Franke & Galkin, 2006).

Die Messungen wurden ungefähr bei der Tonhöhe der mittleren ungespannten Sprechstimmlage begonnen, die vor der Singstimmfeld-Messung für jedes Kind individuell bestimmt wurde, indem ein Sprechstimmfeld mithilfe von DIVAS aufgenommen wurde (Sprechen von Zahlenreihen ab “einundzwanzig” aufwärts). Von dieser mittleren Tonhöhe ausgehend wurden zunächst die tiefen Töne in absteigender Reihenfolge bis zur unteren Stimmgrenze und anschließend die höheren Töne in aufsteigender Reihenfolge bis zur oberen Stimmgrenze gesungen. Nach dem Erfassen aller Töne innerhalb des Stimmumfangs mit der leisesten Stimmgebung folgt derselbe Ablauf für die lautestmögliche Phonation.

Beide Messreihen der minimalen und maximalen Stimmstärke wurden für jeden Ton der C-Dur-Tonleiter innerhalb des individuellen Stimmumfangs jedes Kindes vervollständigt. Bezüglich der höchstmöglichen und tiefstmöglichen Frequenz wurden auch Töne außerhalb der C-Dur-Tonleiter zugelassen, um den Tonhöhenumfang, bzw. seine Grenzen, möglichst exakt zu erfassen. Um allen Probanden zu ermöglichen die vorgegebenen Tonhöhen nachzusingen, war es in manchen Fällen für die Versuchsleiterin nötig den gewünschten Ton einige Male zu wiederholen oder teilweise auch weitere vokalpädagogische Hilfestellungen zu geben, wie z.B. mit einem Glissando von einer bequemen Tonhöhe aus in den jeweiligen Ton zu schleifen o.ä.

Die Versuchsleiterin sang zur Veranschaulichung der Aufgabenstellung Töne mit maximaler und minimaler Stimmintensität vor, um zu gewährleisten,

dass die Begrifflichkeit “so laut bzw. so leise wie möglich zu singen” von allen Probanden richtig erfasst wurde. Es wurden nur stimmliche Modalitäten erfasst, die nicht im Bereich des Schreiens oder Flüsterns lagen (bei letzterem ist eine Grundfrequenzerkennung durch die Stimmfeldsoftware ohnehin nicht möglich). Seidner und Wendler (2004) weisen darauf hin, dass Stimmfeldmessungen somit immer unter physiologischen Gesichtspunkten erfolgen: “So läuft das extrem leise Singen fast ohne Resonanz der Ansatzräume ab und ist nicht einem sängerischen Piano vergleichbar, das deutlich lauter und klangvoller eingesetzt wird. Jedoch ist es auch bei allen sängerischen Fragestellungen stets interessant, wie leicht und wie leise die Kehlkopffunktion anspricht [...]“ (Seidner und Wendler, 2004, S. 208) Identische Messprozeduren wurden zu beiden Erhebungszeitpunkten zu Beginn und zum Ende des Schuljahres (Oktober 2007 bzw. Juni 2008) durchgeführt.

2.2.2 Dysphonia Severity Index (DSI)

Der Dysphonia Severity Index gehört zu den multiparametrischen Methoden der objektiven Erfassung von Stimmqualität und wurde von Wuyts, De Bodt, Molenberghs u.a. (2000) entwickelt. Dieser Index berücksichtigt unterschiedliche Einflüsse auf die Stimmeffizienz und ermöglicht es, den Schweregrad einer etwaigen Stimmstörung zu quantifizieren (Nawka, Franke & Galkin, 2006).

Korrelationen des DSI mit verschiedenen Stimmbeschwerden und Stimmstörungen liegen vor (Kooijman, de Jong, Oudes u.a., 2005) sowie Vergleiche zwischen den DSI-Werten von Studierenden professioneller und nichtprofessioneller Stimmberufe (Timmermans, De Bodt, Wuyts u.a., 2002). Schließlich wurde der DSI herangezogen, um den Ausbildungsverlauf eines Stimmtraining-Programms bei erwachsenen Profisängerinnen und -sängern zu dokumentieren (Timmermans, De Bodt, Wuyts u.a., 2004).

Der DSI wurde auf der Basis einer gewichteten Kombination ausgewählter akustischer Parameter konstruiert, die mit unterschiedlichen Aspekten der Stimmqualität korrelieren: Geringste Stimmintensität (I_{\min}), höchstmögliche Grundfrequenz der Stimme (F_{\max}), maximale Tonhaldedauer (MPT = maximum phonation time) und Jitter (Frequenzstabilität der Stimme).

Die Formel zur Berechnung des DSI ist gegeben durch:

$$\text{DSI} = 0,13 \times \text{MPT} + 0,0053 \times F_{\max} - 0,26 \times I_{\min} - 1,18 \times \text{Jitter} (\%) + 12,4.$$

Der DSI wurde so konstruiert, dass eine normale Stimmqualität mit Werten um +5 korrespondiert und eine stark gestörte Stimme mit Werten um –5

(Wuyts u.a., 2000). Werte die über diesen Zahlenbereich hinausgehen sind ebenso möglich und weisen auf eine besonders gute bzw. extrem schlechte Stimmqualität hin (Hakkesteeft, Brocaar, Wieringa u.a., 2006).

In der vorliegenden Studie wurde der DSI ebenfalls im Rahmen des Stimmdiagnostik-Systems DiVAS gemessen. F_{\max} und I_{\min} werden automatisch aus den Werten der Stimmfeldmessung übernommen. Die maximale Stimmhaltedauer wurde ermittelt, indem alle Probanden aufgefordert wurden den Testvokal [a:] so lange wie möglich auf einer selbstgewählten Tonhöhe bei bequemer Lautstärke zu halten. Die Messbedingungen waren mit denen der Stimmfeldmessungen identisch (normale Raumakustik, Verwendung des XION-Mikrofon-Headsets, Aufnahme in locker stehender Körperhaltung). Die Frequenzstabilität (Jitter) wurde auf Grundlage der Aufnahmen zur maximalen Stimmhaltedauer berechnet.

2.3 Datenauswertung

Der Tonhöhenumfang der Singstimmen kann sowohl aus dem oberen Kurvenzug des Stimmfeldes (Pianokurve = leise Phonation) als auch aus dem unteren Kurvenzug (Fortekurve = laute Phonation) ermittelt werden. Die gemessenen Frequenzwerte wurden dafür in Halbtöne umgerechnet um den Tonhöhenumfang zu erhalten. Somit wurde die Anzahl der Halbtöne innerhalb des Stimmumfangs für jeden Probanden einzeln für die lautestmögliche und die leisestmögliche Stimmgebung zu beiden Messzeitpunkten t_0 (Anfang des Schuljahres) und t_1 (Ende des Schuljahres) ermittelt.

Um den durchschnittlichen Dynamikumfang der einzelnen Probanden zu bestimmen, wurden für jede gemessene Tonhöhe der C-Dur Tonleiter innerhalb des individuellen Tonhöhenumfangs die Werte des Stimmschallpegels (dB) der Pianokurve von denen der Fortekurve subtrahiert. Diese Prozedur wurde nur auf die Frequenzwerte angewendet, die ein Wertepaar von maximaler und minimaler Stimmintensität aufwiesen (die Tonhöhen, bei denen das Kind fähig war den vorgegebenen Ton sowohl sehr laut als auch sehr leise nachzusingen). Die Summe der Differenzwerte wurde anschließend für jeden Probanden durch die Anzahl der erfassten Wertepaare geteilt um den Wert des durchschnittlichen Dynamikumfangs (in dB) zu erhalten. Diese Berechnung erfolgte ebenso für beide Messzeitpunkte t_0 und t_1 .

Zur quantitativen Auswertung der Forschungsfragen wurde das Programm SPSS Version 16.0.2 für Windows verwendet. Um die Auswirkungen des Stimmtrainings in Chorklassen auf Parameter der Stimmgesundheit und Stimmleistung innerhalb eines Schuljahres zu analysieren, wurde ein varianz-

analytisches Design (ANOVA = analysis of variance) mit Messwiederholung im Algorithmus “allgemeines lineares Modell” (ALM) verwendet. Dabei dienten die Variablen “Versuchsgruppe” (Chorklasse und Regelunterricht) und „Geschlecht“ als Zwischensubjektvariablen, während der Messzeitpunkt (t_0 und t_1) als Messwiederholungsvariable eingesetzt wurde. Dieses Auswertungsdesign wurde jeweils einzeln auf die gemessenen abhängigen Variablen (Tonhöhenumfang der Fortekurve und Pianokurve, Dynamikumfang, DSI) angewandt. Für jede abhängige Variable wurde eine einfaktorielle ANOVA durchgeführt um festzustellen, ob signifikante Unterschiede zwischen den Versuchsgruppen zur Basismessung und zum Ende des Schuljahres bestehen.

Um mögliche Auswirkungen des Bildungslevels der Eltern und der Vorerfahrungen im Chorsingen auf die vier abhängigen Messgrößen zu überprüfen, wurden diese Variablen in einer zusätzlichen ANOVA mit Messwiederholung analysiert. Das Alpha-Level betrug bei allen Analysen 0,05.

3 Ergebnisse

3.1 Tonhöhen-, Dynamikumfang und DSI

Tabelle 2 zeigt die deskriptive Statistik zu den abhängigen Variablen beider Messzeitpunkte für jede Gruppe.

		Chorklasse	Regelunterricht
ThU forte	Beginn d. Schuljahrs	20,31 (3,02)	19,22 (3,61)
	Ende d. Schuljahrs	26,56 (3,43)	19,06 (3,10)
ThU piano	Beginn d. Schuljahrs	20,69 (2,87)	20,17 (2,50)
	Ende d. Schuljahrs	25,94 (3,68)	19,17 (2,79)
DyU	Beginn d. Schuljahrs	9,73 (4,07)	7,69 (2,58)
	Ende d. Schuljahrs	15,75 (4,34)	7,38 (2,64)
DSI	Beginn d. Schuljahrs	4,42 (1,16)	4,11 (1,14)
	Ende d. Schuljahrs	6,08 (1,43)	4,22 (1,22)

Tabelle 2: Mittelwerte (und Standardabweichungen) der Gruppen Chorklasse und Regelunterricht für alle abhängigen Variablen beider Messzeitpunkte (Beginn und Ende des Schuljahres)

Hinweis: ThU forte = Tonhöhenumfang bei maximaler Stimmstärke in Halbtönen; ThU piano = Tonhöhenumfang bei minimaler Stimmstärke in Halbtönen; DyU = Dynamikumfang in dB(A); DSI = Dysphonia Severity Index (genauere Erläuterungen siehe Text).

Tabelle 3 fasst die Ergebnisse der ANOVA mit Messwiederholung für alle vier abhängigen Variablen zusammen. Mangels signifikanter Geschlechtereffekte wurde diese Variable von weiteren Analysen ausgeschlossen. Wie die Werte in der Tabelle zeigen, weisen die beobachteten Effekte für alle gemessenen Variablen ähnliche Strukturen auf. Beide Probandengruppen weisen zu Beginn des Schuljahres nicht signifikant unterschiedliche Werte hinsichtlich ihrer stimmlichen Leistungsfähigkeit auf. Die im Laufe des Schuljahres beobachteten Veränderungen deuten auf signifikante Verbesserungen der Stimmleistungen hin, die nur die Chorklassenkinder und nicht die Kinder der Vergleichsgruppe betreffen.

	dF	Zeit	Gruppe x Zeit
		F (partielles η^2)	F (partielles η^2)
ThU forte	(1; 46)	63,51*** (.580)	75,42*** (.621)
ThU piano	(1; 46)	20,82*** (.312)	48,43*** (.513)
DyU	(1; 46)	27,47*** (.374)	33,29*** (.420)
DSI	(1; 46)	23,01*** (.333)	17,78*** (.279)

Tabelle 3: Zusammenfassung der Ergebnisse der ANOVA mit Messwiederholung: F-Werte, Effektstärken, Haupteffekte und Interaktion zwischen Gruppe und Messwiederholungsfaktor für alle vier abhängigen Variablen

Hinweis: ThU forte = Tonhöhenumfang bei maximaler Stimmstärke in Halbtönen; ThU piano = Tonhöhenumfang bei minimaler Stimmstärke in Halbtönen; DyU = Dynamikumfang in dB(A); DSI = Dysphonia Severity Index (genauere Erläuterungen siehe Text). *** $p < .001$

Abbildungen 1 und 2 (siehe Seite 242) zeigen die Mittelwerte der Tonhöhenumfänge für die lautestmögliche und leisestmögliche Stimmgebung beider Messzeitpunkte (t_0 und t_1). Nach der Intervention unterschieden sich die Tonhöhenumfänge der beiden Probandengruppen signifikant voneinander; sowohl für die Fortekurve [$F(1;49) = 59,15$; $p < .001$] als auch für die Pianokurve [$F(1;49) = 45,89$; $p < .001$]. Die Probanden der Chorklasse erreichten eine signifikante Erweiterung ihrer Tonhöhenumfänge durch einen Zuwachs von durchschnittlich 6,25 Halbtönen für die Forte-Kurve bzw. 5,25 Halbtönen für die Pianokurve.

Abbildung 3 (siehe Seite 243) zeigt die durchschnittlichen Dynamikumfang beider Probandengruppen für beide Messzeitpunkte. Auch hier ergab eine einfaktorielle ANOVA, dass die Werte zum Ende des Schuljahres sich für die Probandengruppen aufgrund größerer Dynamikumfang bei den Chorklassenkindern signifikant unterschieden [$F(1;49) = 55,31$; $p < .001$]. Der Dynamikzuwachs in der Gruppe Chorklasse umfasst hierbei durchschnittlich 6,02 dB.

Abbildung 4 (siehe Seite 243) zeigt den Interaktionseffekt zwischen Gruppe und Zeit für die Durchschnittswerte des DSI. Zum Ende des Schuljahres (t_1) zeigen sich auch bezüglich der durch den DSI erfassten Stimmqualität signifikante Unterschiede zwischen den Probandengruppen [$F(1;49) = 21,66$; $p < .001$]. Kinder in der Chorklasse verbesserten ihre DSI-Werte um durch-

schnittlich 1,66 Punkte, während Kinder in der Vergleichsgruppe keine signifikante Verbesserung erreichten.

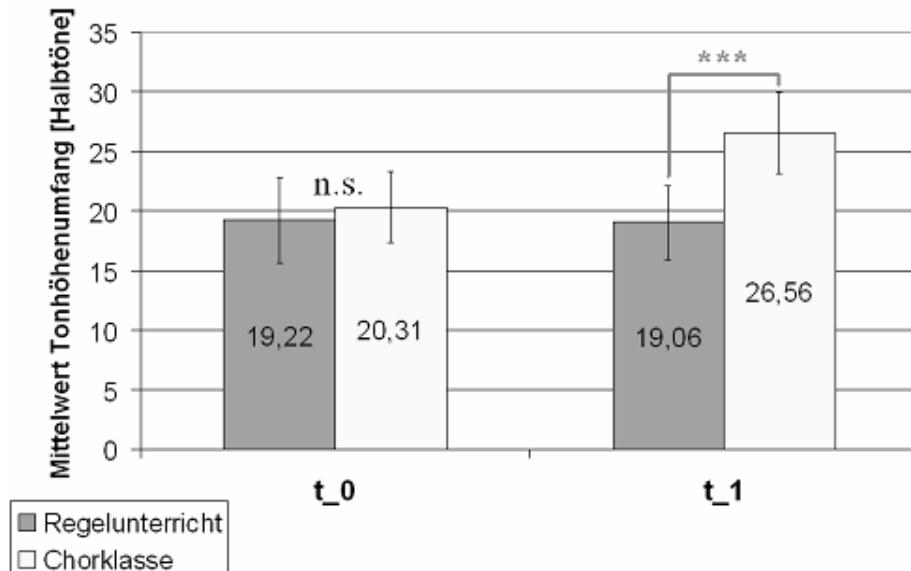


Abb. 1: Mittelwerte (und Standardabweichungen) des Tonhöhenumfangs in Halbtönen für die Fortekurve (lauteste Stimmgebung) des Stimmfeldes beider Probandengruppen zu Beginn des Schuljahres (t0) und zum Ende des Schuljahres (t1).

*** Statistisch signifikant bei $p < .001$

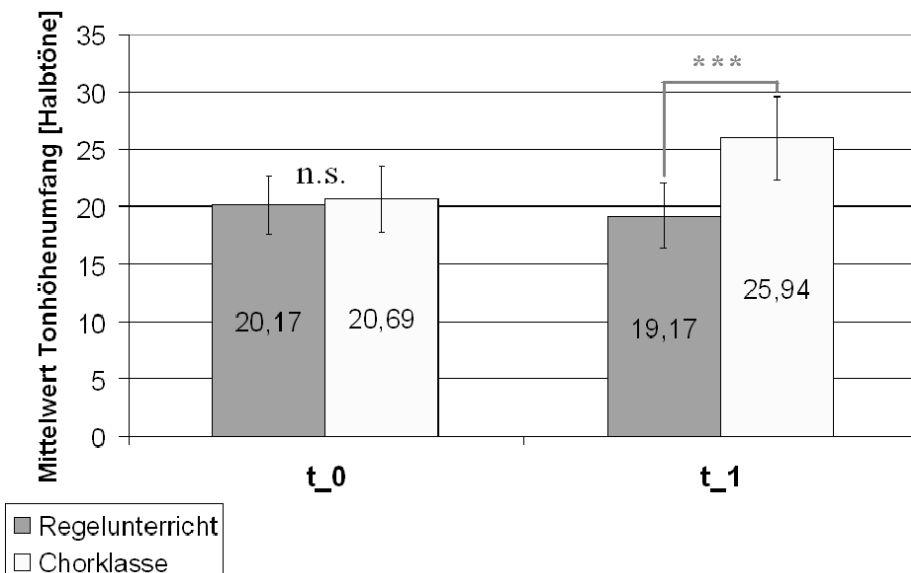


Abb. 2: Mittelwerte (und Standardabweichungen) des Tonhöhenumfangs in Halbtönen für die Pianokurve (leiseste Stimmgebung) des Stimmfeldes beider Probandengruppen zu Beginn des Schuljahres (t0) und zum Ende des Schuljahres (t1).

*** Statistisch signifikant bei $p < .001$

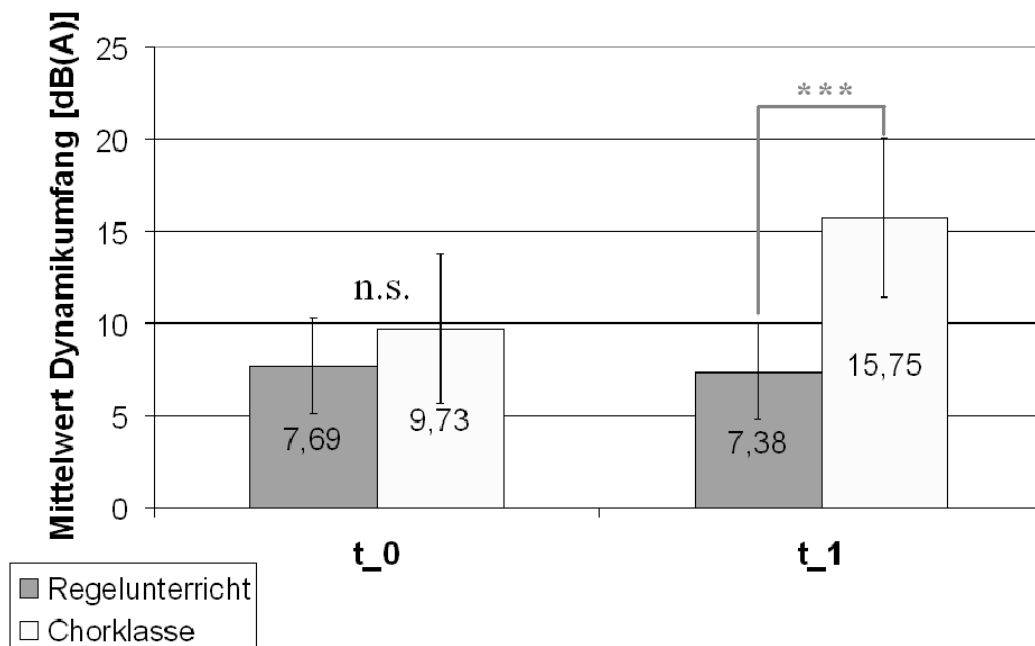


Abb. 3: Mittelwerte (und Standardabweichungen) des Dynamikumfangs des Stimmfeldes in dB (A) beider Probandengruppen zu Beginn des Schuljahres (t0) und zum Ende des Schuljahres (t1). *** Statistisch signifikant bei $p < .001$

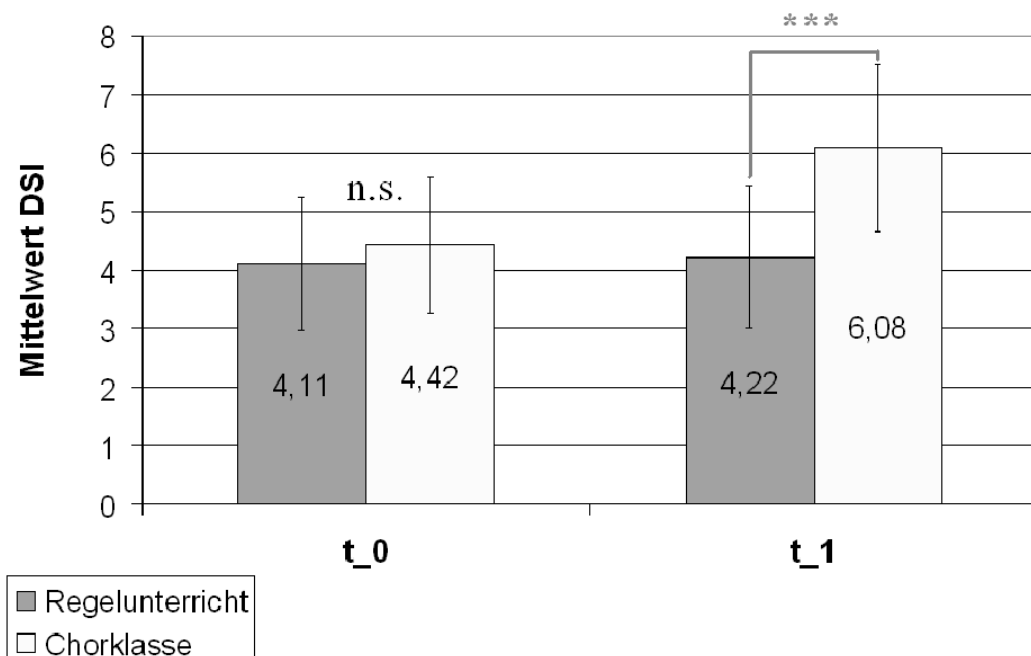


Abb. 4: Mittelwerte (und Standardabweichungen) der DSI-Werte für beide Probandengruppen zu Beginn des Schuljahres (t0) und zum Ende des Schuljahres (t1). *** Statistisch signifikant bei $p < .001$

3.2 *Bildungslevel der Eltern und Vorerfahrung im Chorsingen*

Das Bildungslevel der Eltern verursachte keine signifikanten Haupteffekte bezüglich des Tonhöhen- und Dynamikumfangs. Es konnte jedoch ein signifikanter Zwischensubjekteffekt bezüglich des DSI festgestellt werden [$F(1;45) = 4,69$; $p < .036$; partielles $\eta^2 = .094$]. Die Probanden mit Eltern höherer Bildungslevel wiesen durchschnittlich bessere Werte dieses Indexes zur Stimmgesundheit und Stimmqualität auf. Es konnten weiterhin keine Interaktionen zwischen dem Bildungslevel und den Faktoren Gruppe und Zeit für alle gemessenen Größen festgestellt werden. Diese Analysen zeigen, dass der Bildungsstand der Eltern im Rahmen der vorliegenden Untersuchungen keinen Einfluss auf die längsschnittliche Verbesserung der abhängigen Variablen Tonhöhen-, Dynamikumfang und DSI hatte.

Probanden mit Vorerfahrung im Chorsingen unterschieden sich systematisch von denen ohne Chorerfahrung. Dies führte zu signifikanten Haupteffekten dieser Variable bezüglich des Tonhöhenumfangs (Pianokurve: [$F(1;46) = 3,01$; $p < .089$; partielles $\eta^2 = .061$]; Fortekurve: [$F(1;46) = 7,14$; $p < .010$; partielles $\eta^2 = .134$]) und des Dynamikumfangs [$F(1;46) = 6,98$; $p < .011$; partielles $\eta^2 = .132$]. Es konnten jedoch weder Effekte der Chorerfahrung bezüglich des DSI beobachtet werden, noch gab es Interaktionen der Vorerfahrung in Chören mit den Faktoren Gruppe und Zeit. Dies bedeutet, dass das vorherige Singen in Chören keinen Einfluss auf die zeitliche Verbesserung der vier gemessenen Variablen bei den Chorklassenkindern hatte.

4 **Diskussion**

Die vorliegende Längsschnittstudie ist Auswirkungen von Chorklassenunterricht auf die Entwicklung der Singstimme von Kindern an einer Gesamtschule gewidmet. Insbesondere wurde eruiert, inwiefern Stimmtraining in Chorklassen Einfluss auf Stimmleistung, Stimmqualität und Stimmgesundheit von Fünftklässlern hatte.

Es wurden signifikante Unterschiede zwischen den beiden Probandengruppen in Bezug auf medizinisch-phoniatrische Indikatoren der Stimmleistung, Stimmqualität und Stimmgesundheit festgestellt. Diese Beobachtungen bestätigen die Annahme, dass Stimmtraining in Chorklassen langfristig die Beherrschung des Stimmapparates begünstigen und darüber hinaus zur Prävention von Stimmstörungen beitragen kann. Chorklassen fördern damit nicht allein musikalische Potentiale, sondern versetzen Kinder in die Lage, mit ihren Stimmwerkzeugen ökonomisch und physiologisch angemessen umgehen zu

können. Ansätze zur Integration eines „Stimmheilunterrichts“ in den schulischen Musikunterricht werden beispielsweise auch von Sallat (2001) aufgezeigt, der Möglichkeiten erläutert, eine Behandlung von gestörten Stimmen mit den Inhalten des Lehrplanes zu vereinbaren. Beobachtungen von Sallat (2009) über Zusammenhänge musikalischer und sprachlicher Fähigkeiten bei sprachgestörten im Vergleich zu sprachlich normal entwickelten Kindern können somit möglicherweise hinsichtlich positiver Einflüsse von Gesangstraining auf vokale Fähigkeiten und Fertigkeiten gesunder Kinder ausgeweitet werden.

Bemerkenswert erscheint, dass Nicht-Chorklassenkinder nicht nur keine bedeutenden Fortschritte in ihrer stimmlichen Entwicklung anhand der hier gemessenen Indikatoren erzielen konnten, sondern sogar in ihrer Entwicklung stagnierten. Schon aus diesem Grund erscheinen weitere Untersuchungen notwendig, die noch größere Zeitfenster einschließen.

Obwohl die beobachteten Effekte für alle gemessenen Parameter dieselbe Struktur aufwiesen und sich sogar auf die Verbesserung eines klinischen Index zur Stimmgesundheit, des Dysphonia Severity Index, erstreckten, sollen im Folgenden die unterschiedlichen Implikationen dieser Ergebnisse für alle Parameter einzeln diskutiert werden

4.1 Tonhöhenumfang

Die Ergebnisse für die lauteste und leiseste Stimmgebung der Stimmfeldmessung verdeutlichen, dass Chorklassenkinder in Folge des speziellen Stimmtrainings ihren Tonhöhenumfang deutlich erweitern konnten, wohingegen bei Kindern im Regelunterricht die Grenzen des Tonhöhenumfangs der Singstimme stagnierten.

Die signifikante Erweiterung des Tonhöhenumfangs der Chorklassenkinder kann durch eine effektivere Nutzung und Steuerung der laryngealen Muskulatur und eine bessere Kontrolle über das Atemsystem erklärt werden. Ein optimales Zusammenspiel von Atmung und Stimmgebung befähigt zu einer flexibleren Feinabstimmung der physiologischen Mechanismen des Phonationsapparates und beugt somit einem stimmlichen Fehlverhalten vor. Einseitiges Beanspruchen eingeschränkter Bereiche des Stimmumfangs, beispielsweise aufgrund mangelnder Beherrschung der unterschiedlichen Registerfunktionen, geht mit einer einseitigen Muskelaktivität her. Dies führt häufig zu Fehlspannungen, die in Stimmstörungen resultieren können. Erweiterungen des Stimmumfangs, die einen Ausbau der Brust- und Kopfstimmfunktion beinhalten, um-

fassen aus stimmgesundheitslicher Sicht also auch einen hohen stimmprophylaktischen Wert.

Die Beobachtung, dass der Tonhöhenumfang für die laute Stimmgebung stärker zunahm als für die leise (6,25 vs. 5,25 Halbtöne) kann plausibel gemacht werden, indem der allgemeine Erfahrungswert von Gesangspädagogen und Sängern herangezogen wird, dass die kontrollierte Stimmgebung im Piano und Pianissimo eine der schwierigsten Herausforderungen der Gesangsstimme ist. "The ability of a singer to maintain minimal intensity while maintaining frequency requires exquisite coordination of respiration, phonation and resonance." (Le Borgne & Weinrich, 2002). Aus dieser Sicht ist es umso beachtlicher, dass die Erweiterung des Tonhöhenumfangs für die leisestmöglichen Töne dennoch durchschnittlich mehr als 5 Halbtöne betrug. Leise Stimmgebung zu beherrschen, ist aus stimmgesundheitslicher Sicht eine Grundvoraussetzung, um stimmliche Belastungen und Überanstrengungen zu vermeiden, die sonst dauerhaft zu funktionelle Dysphonien (Störungen der Sprechstimme) und Dysodien (Störungen der Singstimme) führen können.

Generell scheint ein Zuwachs von durchschnittlich fünf bis sechs Halbtönen Tonhöhenumfang innerhalb eines Schuljahres beachtlich; insbesondere wenn man diese Ergebnisse mit Längsschnittstudien von Erwachsenen vergleicht. Le Borgne & Weinrich (2002), beispielsweise, berichten von einem Zuwachs von 1,38 Halbtönen bei Gesangsstudierenden nach einer neunmonatigen intensiven Trainingsphase. Mendes u.a. (2003) konnten für Gesangsstudierende innerhalb von vier Semestern eine Erweiterung des Tonhöhenumfangs um 5,7 Halbtöne zeigen.

Der stagnierende Tonhöhenumfang der Kinder im Regelunterricht Musik kann hingegen durch die fehlende Fähigkeit, die volle Kapazität des Stimmapparates auszunutzen, erklärt werden. Es kann vermutet werden, dass Ursachen auch in einer ungenügenden Förderung der kindlichen Singstimme in Kindergarten und Grundschule liegen (vgl. dazu empirische Ergebnisse von Brünger, 2003).

Die in der vorliegenden Studie beobachteten Tonhöhenumfänge zum Ende des Schuljahres (Gruppe Chorklasse: 19,06 Halbtöne, Gruppe Regelunterricht: 26,56 Halbtöne) stimmen überein mit aktuellen Literaturangaben bezüglich querschnittlicher Untersuchungen von sängerisch aktiven und nicht-aktiven Kindern. Fuchs u.a. (2006) berichten durchschnittliche Tonhöhenumfänge von 18,9 Halbtönen für nichtsingende Jungen (Alter 11-13,5 Jahre) und 22,3 Halbtönen für nichtsingende Mädchen derselben Altersgruppe. Innerhalb derselben Querschnittsuntersuchung beobachteten die Autoren einen Tonhöhenumfang

von 26,5 Halbtönen für sängerisch aktive Mädchen und Jungen dieser Altersgruppe.

4.2 *Dynamikumfang*

Die Studie zeigt, dass sich die Kontrolle der Stimmdynamik bei den Chorklassenkindern verbessert. Diese Kinder erreichen einen Dynamikzuwachs von durchschnittlich 6,02 dB, während der Dynamikumfang der Kinder im Regelunterricht nahezu konstant blieb. Hierbei ist zu beachten, dass ein Unterschied von etwa 10 dB einer Verdopplung der wahrgenommenen Lautstärke entspricht (Terhardt, 1998). In Anbetracht der Ergebnisse vorliegender Längsschnittstudien an erwachsenen Gesangsstudierenden, die im Verlauf von vier bis fünf Jahren Ausbildungsdauer ihre maximale Stimmintensität um durchschnittlich (gemittelt über Probanden, Vokale und Tonhöhen) 2,4 dB erweitern konnten (Mürbe, Sundberg, Iwarsson u.a., 1999), erscheint ein Zuwachs von 6 dB innerhalb eines Schuljahres als ein enormer Fortschritt.

Die Fähigkeit die Singstimme mit einer signifikant verbesserten dynamischen Flexibilität zu steuern, kann im Zusammenhang mit Stimmtraining in Chorklassen durch eine verbesserte Koordination der Atmung und Stimmgebung interpretiert werden, die sich vermutlich vor allem durch ein verbessertes Zusammenspiel von subglottischem Druck und Stimmlippenspannung ausdrückt. Eine physiologisch erzeugte Lautstärkeregulierung der Stimme resultiert automatisch in einem ökonomischeren Stimmgebrauch; ein Hauptelement von Stimmhygiene und der Prophylaxe von Stimmstörungen.

Es bleibt zu vermuten, dass eine kontrollierte Steuerung der Stimmintensität und eine Erweiterung des Dynamikumfangs beim Singen sich auch positiv auf die Lautstärkesteuerung der Sprechstimme auswirken. Eine tragfähige und belastbare Stimme ist als Grundvoraussetzung aller Prozesse sprachlichen Handelns anzusehen und bildet somit die Basis einer erfolgreichen stimmlichen Kommunikation. Sallat (2009) beschreibt den Gebrauch der Sprechwerkzeuge als eine wesentliche Schlüsselkompetenz, die einen zentralen Platz in Bildungs- und Erziehungsplänen aufweisen sollte.

4.3 *Dysphonia Severity Index (DSI)*

Ein signifikanter Anstieg des DSI, der nur für die Kinder der Chorklasse festgestellt werden konnte, impliziert eine Verbesserung der Stimmqualität und Stimmgesundheit, die auf das Stimmtraining zurückzuführen ist. Eine Reihe von Studien belegt den Zusammenhang von DSI-Werten und objektiven Parametern der Stimmqualität und Stimmgesundheit. Der DSI korreliert u.a. mit

dem “Voice Handicap Index” zur Bewertung von Stimmstörungen (Wuyts u.a., 2000) mit Heiserkeitsbewertungen anhand der GRBAS-Skala und dem Schweregrad von Stimmstörungen (Hakkesteeft, Brocaar, Wieringa u.a., 2008a). Je schlechter diese Parameter der Stimmqualität und Stimmgesundheit waren, desto niedriger wurden auch die Werte des DSI. Ebenso zeigten Studien signifikante Verbesserungen des DSI bei stimmkranken Patienten nach erfolgreichen Stimmbehandlungen (Evans und Nawka, 2008, Hakkesteeft, im Druck).

Somit sind die vorliegenden Ergebnisse von besonderer Relevanz, da sie eine positive Auswirkung des Stimmtrainings über musikalisch-ästhetische Ziele hinaus implizieren. Eine signifikante Verbesserung dieses klinischen Stimmgesundheits-Indexes verweist auf präventive Implikationen des Stimmtrainings bezüglich der kindlichen Stimmhygiene und -gesundheit. Diese Interpretation wird zum Teil dadurch gestützt, dass nach Aussage der Eltern keines der an der Studie teilnehmenden Kinder Stimmkrankheiten aufwies und die gemessenen Stimmleistungsparameter keine klinischen Stimmprobleme vermuten lassen.

Nawka u.a. (2006) beschreiben, dass eine normale Stimmqualität mit DSI-Werten > 4.2 korreliert. Von dieser Angabe ausgehend scheinen alle Probanden der vorliegenden Studie im Bereich normaler und gesunder Stimmen zu sein. Unserem Wissen nach existieren bisher keine Normwerte des DSI für Kinder, sondern nur für Erwachsene (Hakkesteeft u.a., 2006). Die Test-Retest-Variabilität des DSI wurde bisher ebenfalls nur für Erwachsene untersucht (Hakkesteeft, Wieringa, Brocaar u.a., 2008). Hier zeigen die Ergebnisse, dass die Werte des DSI dann signifikant werden, wenn Unterschiede innerhalb eines Probanden größer als 2,49 sind. Diese Information scheint eine Einschränkung der vorliegenden Ergebnisse darzustellen, da hier bereits Unterschiede von 1,66 statistisch signifikant wurden. Da aber andererseits in Untersuchungen an Erwachsenen gezeigt wurde, dass der DSI eine altersabhängige Größe ist (Hakkesteeft u.a., 2006) und vergleichbare Erhebungen an Kindern bisher nicht durchgeführt wurden, lassen sich diese Angaben nicht oder nur eingeschränkt auf die vorliegenden Erhebungen übertragen. Aufgrund der identischen Struktur der Effekte für alle gemessenen Stimmleistungsparameter muss die Signifikanz der vorliegenden Ergebnisse anerkannt werden.

4.4 Einflüsse des Umfeldes

In den vorliegenden Erhebungen wiesen die Kinder mit höherem Bildungslevel der Eltern bessere Absolutwerte des DSI auf. Dies weist auf einen Einfluss

des sozio-ökonomischen Status' der Eltern auf die individuellen Ressourcen der Stimmgesundheit hin. Da der Anstieg der untersuchten Parameter nicht mit dem Bildungslevel der Eltern korrelierte, hatte dieser systematische Unterschied zwischen den Kindern keinen Einfluss auf die Verbesserung der Stimmleistungen und Stimmqualität. Diese Ergebnisse lassen vermuten, dass vokalpädagogische Förderprogramme an Schulen nicht nur Ressourcen der Stimmgesundheit im Sinne einer Prävention verfügbar machen, sondern dass diese Förderung unabhängig von sozio-ökonomischen Einflüssen des Elternhauses alle Kinder profitieren lässt. Im Licht der Ergebnisse der PISA-Studien, die einen Zusammenhang zwischen Sprachkompetenz und Bildungschancen bei Kindern und Jugendlichen zeigten, erhalten diese Ergebnisse besondere Relevanz.

Den Erwartungen entsprechend wiesen die Kinder mit Vorerfahrungen im Chorsingen größere Absolutwerte der gemessenen Stimmleistungsparameter auf (Tonhöhenumfang der lauten und leisen Stimmgebung und Dynamikumfang). Da der Zugewinn der Stimmleistungsparameter jedoch unabhängig von der vorherigen Teilnahme an Chören war, scheint die Verbesserung der Stimmleistungen in Chorklassen unabhängig von den Vorerfahrungen zu sein.

4.5 Einschränkungen der Studie

Die Generalisierbarkeit der vorliegenden Ergebnisse bleibt aufgrund der Stichprobengröße und der teilweisen Selbstselektion der Versuchs- und Vergleichsgruppe offen. Obwohl die Teilnahme an der Chorklasse durch ein schulinternes Verfahren per Los entschieden wurde und kein Stichprobenschwund innerhalb des Erhebungszeitraums von einem Schuljahr eintrat, muss berücksichtigt werden, dass die Chorklassenkinder unter einer Gruppe von Bewerbern ausgelost wurden, deren Eltern sie für dieses Förderprogramm angemeldet hatten. Diese Auswahl entspricht nicht in allen Fällen den Wünschen der Kinder. 21,9 % der Chorklassenkinder geben an, dass es nicht ihr eigener Wunsch war an diesem Förderprogramm teilzunehmen.

Die externe Validität dieser explorativen Ergebnisse bleibt auf die speziellen Probandengruppen in ihren Kontextfaktoren beschränkt. Hier sind weitere Studien nötig um den Einfluss des Umfelds, wie z. B. des sozialen Umfelds, zu eruieren. Künftige Studien müssten differenzieren, ob die Fortschritte ein Effekt von speziellen Faktoren des Unterrichts waren, wie beispielsweise bestimmter Methoden der Stimmbildung. Es bleibt offen, ob die Erfolge die in dieser Studie gemessen wurden von einzelnen oder multidimensionalen Aspekten der Intervention bewirkt wurden. Hier sind weitere Studien nötig, die

den Einfluss der einzelnen Interventionen im komplexen Wirkungsgefüge näher beleuchten. Schließlich ist die Aussagekraft der Ergebnisse auf die Altersgruppe dieser Probanden beschränkt und wirft die Frage nach einem günstigen Zeitpunkt für Stimmförderung von Kindern auf, die vermutlich idealerweise bereits im Kindergarten und in der Grundschule beginnen sollte.

Danksagungen

Das vorliegende Forschungsprojekt wurde durch ein Promotionsstipendium der Studienstiftung des Deutschen Volkes ermöglicht.

Ein großer Dank gilt der Medizinelektronikfirma XION Medical GmbH Berlin, die großzügigerweise die Messausrüstung und Software DiVAS für die Stimmfeldmessungen zur Verfügung stellte. Des Weiteren gilt ein großer Dank den zahlreichen Lehrerinnen und Lehrern der Helene-Lange-Schule in Oldenburg, die an der Organisation der Erhebungen beteiligt waren, sowie den Schülerinnen und Schülern die an den Untersuchungen teilgenommen haben.

Literatur

- Åkerlund L. & Gramming, P. (1994). Average loudness level, mean fundamental frequency, and subglottal pressure comparison between female singers and non-singers. *Journal of Voice* 8(3), 263-270.
- Awan S. (1991). Phonetographic profiles and F0-SPL characteristics of untrained versus trained vocal groups. *Journal of Voice* 5(1), 41-50.
- Bailey, B. & Davidson, J. W. (2003). Amateur group singing as a therapeutic agent. *Nordic Journal of Music Therapy* 12, 18-32.
- Berger, R. & Walde, V. (2002). Zur Stimmentwicklung bei Grundschulkindern und ihren musischen Aktivitäten. *Aktuelle phoniatisch-pädaudiologische Aspekte* 10, 111-116.
- Böhme G., & Stuchlik G. (1995). Voice profiles and standard voice profiles of untrained children. *Journal of Voice* 9(3), 304-307.
- Brünger, P. (2003). Singen im Kindergarten. Eine Untersuchung unter bayerischen und niedersächsischen Kindergartenfachkräften. Augsburg: Wißner (Forum Musikpädagogik Band 56).
- Clift, S. & Hancox, G. (2001). The perceived benefits of singing: Findings from preliminary surveys with a university college choral society. *Journal of the Royal Society for the Promotion of Health* 121, 4, 248-256.
- Coleman R. (1993). Sources of variation in phonetograms. *Journal of Voice* 7(1), 1-14.

- Connor, N. P., Cohen, S. B., Theis, S. M., Thibeault, S. L., Heatley, D. G. & Bless, D. M. (2008). Attitudes of children with dysphonia. *Journal of Voice* 22(2), 197-209.
- Dejonckere, P. H., Wieneke, G. H., Bloemenkamp, D., Lebacqz, J. (1996). Fo-perturbation and Fo/loudness dynamics in voices of normal children, with and without education in singing. *International journal of pediatric otorhinolaryngology* 35(2), 107-15.
- Evans, R. & Nawka, T. (2008). Evaluation der Stimmtherapie von Patienten mit funktionellen Dysphonien. Deutsche Gesellschaft für Phoniatrie und Pädaudiologie. 25. Wissenschaftliche Jahrestagung der Deutschen Gesellschaft für Phoniatrie und Pädaudiologie. Düsseldorf, 12.-14.09.2008. Düsseldorf: German Medical Science GMS Publishing House; 2008. Doc08dgppV55
- Fuchs, M., Heide, S., Hentschel, B., Gelbrich, G., Makuch, A., Thiel, S., Täschner, R., Dietz, A. (2006). Vocal efficiency parameters in children and adolescents: effect of physical development and singing activity. *HNO* 54(12), 971-80.
- Gramming, P. (1988): The phonetogram, an experimental and clinical study. Dissertation, Lund University.
- Hakkesteeft, M. M., Brocaar, M. P., Wieringa, M. H., Feenstra, L. (2006). Influence of age and gender on the Dysphonia Severity Index. *Folia Phoniatrica et Logopaedica* 58, 264-273.
- Hakkesteeft, M. M., Wieringa, M. H., Brocaar, M. P., Mulder, P. G., Feenstra, L. (2008). The interobserver and test-retest variability of the dysphonia severity index. *Folia Phoniatrica et Logopaedica* 60, 86-90.
- Hakkesteeft, M. M., Brocaar, M. P., Wieringa, M. H., Feenstra, L. (2008a). The relationship between perceptual evaluation and objective multiparametric evaluation of dysphonia severity. *Journal of Voice* 22(2), 138-145.
- Hakkesteeft, M. M., Brocaar, M. P., Wieringa, M. H. (im Druck). The applicability of the Dysphonia Severity Index and the Voice Handicap Index in evaluating effects of voice therapy and phonosurgery. *Journal of Voice*, in press.
- Hoos de Jokisch, B. (2003). Verlust der Stimme - Verlust des Körpers? *Üben und Musizieren* 3, 6-13.
- Jentschke, S., Koelsch, S., Sallat, S., Friederici, A. D. (2008). Children with specific language impairment also show impairment of music-syntactic processing. *Journal of Cognitive Neuroscience* 20(11), 1940-1951.

- Kooijman, P. G., de Jong, F. I., Oudes, M. J., Huinck, W., van Acht, H., Graamans, K. (2005). Muscular tension and body posture in relation to voice handicap and voice quality in teachers with persistent voice complaints. *Folia Phoniatica et Logopaedica* 57, 134–147.
- Kreutz, G. (2004). Wirkungen und Bedeutungen des Chorsingens: Psychophysiologische und sozialpsychologische Aspekte. In G. Kreutz & J. Bähr (Eds.) *Anstöße – Musikalische Bildung fordern und fördern*. Festschrift Hans Günther Bastian zum 60. Geburtstag (pp. 107-124). Augsburg: Wißner.
- Kreutz, G., Bongard, S., Grebe, D., Rohrmann, S. & Hodapp, V. (2004). Effects of choir singing or listening on secretory IgA, cortisol, and emotional state. *Journal of Behavioral Medicine* 27(6), 623-634.
- Lass, N. J., Ruscello, D.M., Stout, L. L., Hoffmann, F. M. (1991). Peer perceptions of normal and voice-disordered children. *Folia Phoniatica et Logopaedica* 43, 29–35.
- Le Borgne, W. D. & Weinreich, B. D. (2002). Phonetogram changes for trained singers over a nine-month period of vocal training. *Journal of Voice* 16(1), 37-43.
- Mendes, A. P., Rothman, H. B., Sapienza C., Brown W. S. (2003). Effects of vocal training on the acoustic parameters of the singing voice. *Journal of voice* 17(4), 529-43.
- Mürbe, D., Sundberg, J., Iwarsson, J., Pabst, F., Hofmann, G. (1999). Longitudinal study of solo singer education effects on maximum SPL and level in the singers' formant range. *Logopedics Phoniatics Vocology* 24, 178-86.
- Nawka, T., Franke, I., Galkin, E. (2006). Objektive Messverfahren in der Stimmdiagnostik. *Forum Logopädie* 4 (20), 14-21.
- Ruscello D.M., Lass N.J., Podbesek J. (1988). Listeners' perceptions of normal and voice-disordered children. *Folia Phoniatica et Logopaedica* 40, 290–296.
- Sallat, S. (2009): *Der Ton macht die Musik - und die Sprache*. *Logos Interdisziplinär* 17 (2), 84-92.
- Sallat, S. (2001). *Störungen der Stimme und ihre ganzheitliche Behandlung im Musikunterricht*.
Online im Internet: URL: <http://www.foepaed.net/sallat/stimme.pdf>. (Zugriff am 22.7.2009)
- Seidner, W. & Wendler, J. (2004). *Die Sängerstimme: Phoniatische Grundlagen des Gesangs*. Berlin: Henschel.

- Schutte, H. & Seidner, W. (1983). Recommendation by the Union of European Phoniaticians: Standardizing voice area measurement – Phonetography. *Folia Phoniatica et Logopaedica* 35, 286-288.
- Sulter, A. M., Schutte, H. K., Miller, D. G. (1995). Differences in phonetogram features between male and female subjects with and without vocal training. *Journal of Voice* 9(4), 363-377.
- Terhardt, E. (1998). *Akustische Kommunikation* (p. 279) Berlin: Springer-Verlag.
- Ternström, S. (1991). Physical and acoustic factors that interact with the singer to produce the choral sound. *Journal of Voice* 5(2), 128-143.
- Timmermans, B., De Bodt, M. S., Wuyts, F. L., Boudewijns, A., Clement, G., Peeters, A., Van de Heyning, P. (2002). Poor voice quality in future elite vocal performers and professional voice users. *Journal of Voice* 16(3), 372–382.
- Timmermans, B., De Bodt, M. S., Wuyts F. L., Van de Heyning, P. (2004). Training outcome in future professional voice users after 18 months of voice training. *Folia Phoniatica et Logopaedica* 56, 120–129.
- Von Gutzeit, R. (2001). Singe, wem Gesang gegeben. *Üben und Musizieren* 2, 1.
- Wuyts, F.L., Heyler L., Mertens F., De Bodt M., Van de Heyning P. H. (2002). Normative voice range profiles of untrained boys and girls. *Journal of Voice* 16(4), 460–465.
- Wuyts, F., De Bodt, M., Molenberghs, G., Remacle, M., Heylen, L., Millet, B., Van Lierde, C., Raes, J., Van de Heyning, P. (2000). The dysphonia severity index: An objective measure of vocal quality based on a multiparameter approach. *Journal of Speech, Language and Hearing Research* 43(3), 796–809.
- Züchner, I., Arnoldt, B., Vossler, A. (2008). Kinder und Jugendliche in Ganztagsangeboten. In Holtappels, H.G., Klieme, E., Rauschenbach, T., Stecher, L. (Eds.) *Ganztagsschule in Deutschland. Ergebnisse der Ausgangserhebung der "Studie zur Entwicklung von Ganztagschulen" (StEG). Studien zur ganztägigen Bildung.* (pp. 106-123) Band 1. Weinheim: Juventa.

Musiktheorie in musizierpraktischem Schulunterricht

Zur Effizienz basaler Harmonielehre für das Improvisieren mit Keyboards

1 Einleitung

Die Diskussion um die Sinnhaftigkeit musiktheoretischer Inhalte im schulischen Musikunterricht wird seit einem halben Jahrhundert unter spürbarem Zwang didaktisch-konzeptioneller Fraktionen geführt. Auch nach dem Siegeszug handlungsorientierten Unterrichts hat sich an der Polarisierung der Positionen bis heute kaum etwas geändert: Für die einen ist Musiktheorie (hier verstanden als Material und Struktur von Musik) das unverzichtbare musikalische „ABC“, andere warnen noch immer vor *grammatischem Fundamentalismus* (Ehrenforth, zit. bei Waldenfels 2001, S. 23) oder zweifeln die Notwendigkeit musikalischer Elementarlehre in allgemein bildenden Schulen generell an (einschlägige Literatur zu dieser Debatte siehe Kühn 2002, Gruhn 2003, Bäßler/Nimczik 2005, Mazurowicz 2005, Froebe 2006, Hempel 2006, Richter 2007, Lang 2008). Davon offenbar unberührt, fristet Musiktheorie im Schulalltag ein Schattendasein zwischen Lehrplanverpflichtung und mangelhaftem methodischen Fundament. Zur spärlichen Beliebtheit bei Lehrern wie auch Schülern mag beitragen, dass wenig darüber bekannt ist, wie sich die Lehre musikalischer Strukturen konkret auf die Fähigkeiten und Fertigkeiten der Schüler auswirken kann. Dabei hat die Musikpsychologie hier schon Beachtliches geleistet, das der didaktisch-methodischen Aufbereitung harrt, vor allem hinsichtlich des Einflusses von theoretischem Wissen auf musikbezogene Gedächtnisleistung (z. B. Boltz 1991, Leman 2000, im Überblick bei De la Motte-Haber 2005 und Bruhn 2005).

Weitaus weniger Erkenntnisse liegen zur Schnittstelle von Theorie und generativer musikalischer Performanz vor (vgl. Lehmann 2005). Welche neuen Möglichkeiten ergeben sich aus neu erworbenem strukturellem Wissen für eine Improvisation?

Antworten hierauf dürften sich für die Musikpädagogik als echtes Desiderat erweisen, hat doch das Improvisieren seit seiner Funktionalisierung als „Vorbereitung auf das Werkhören“ durch M. Alt (zit. bei Noll 1973, 76) kontinuierlich an Bedeutung für das gesamte Spektrum möglicher Unterrichtsaktivitäten und -inhalte gewonnen. Trotzdem taucht in Schulbüchern und anderen Arbeitsmaterialien Improvisation im Kontext der Theorie nur ausnahmsweise auf. Fantasiervolle Musik-Lehrbücher mit hohen Anteilen von Improvisation (z. B. Vella 2003) meiden Musiktheorie und verbannen sie in tabellarische Anhänge. Progressive Musiktheoriebücher (z. B. Ringel 2009) beschränken sich in puncto Improvisation wiederum auf wenige bewährte Methoden wie das freie Spiel mit Pentatonik, das doch bereits im Orffschen Schulwerk angeboten worden war. Überdies mangelt es in fast sämtlichen Arbeitsmaterialien an Hilfen für ein planmäßiges Auswerten der Arbeitsergebnisse (Schülerimprovisationen).

Ausgangspunkt der hier dokumentierten Studie ist die Auffassung, dass die Fähigkeit zum Improvisieren in modernen Musikstilen darauf beruht, „die zur Verfügung stehenden musikalischen Muster und Formeln aus dem Stegreif zu einer stimmigen und regelbasierten Struktur zusammenzusetzen und zu variieren“ (Gembris 2008, 169 f.). Ebendieses Erlernen einer regelbasierten Struktur – hier aufgefasst als Aspekt von Musiktheorie – wird genauer untersucht. Es handelt sich somit um eine Effizienzstudie zum Lernprozess generativer musikalischer Performanz.

2 Die Studie

Hypothese

Der Erwerb musiktheoretischen Wissens eröffnet für das Improvisieren zusätzliches Lernpotenzial, das schon kurzfristig nachweisbar ist.

Obwohl nahe liegt, dass sich musiktheoretisches Wissen durch seine Nähe zu Reglementierungen zunächst restriktiv auf das Improvisieren auswirkt, wird auch a) das Bewahren und b) das Ausweiten persönlicher Spielweisen beim harmonisch gebundenen Spiel überprüft.

Versuchspersonen, Lerngegenstand, Medien

Für die Studie werden Versuchspersonen benötigt, die vor Beginn des Experiments noch keinerlei Vorwissen auf dem Gebiet der Harmonielehre haben, andererseits aber in einem Alter sind, das Erfahrungen mit Strukturierungsvorgängen und dem Umgang mit Fachbegriffen in anderen Schulfächern mit sich bringt. 26 Schüler der 7. Klasse einer Realschule in Freiberg am Neckar erfüllen diese Kriterien.

Lerngegenstand im engeren Sinne ist der hiermit erst eingeführte tonale Dreiklang (ohne Differenzierung von Dur und Moll). Die Erläuterung dieser basalen harmonischen Struktur wird zur entscheidenden didaktischen Intervention zwischen den Improvisationsphasen. Die Improvisationen sollen sich akustisch an der Situation des Klassenmusizierens orientieren: Sie werden zum Playback des von der Schulklasse selbst zuvor aufgenommenen Shantys *What shall we do with the drunken sailor* ausgeführt, dem bekanntermaßen nur zwei Akkorde zugrunde liegen (z.B. d-Moll/C-Dur).

What shall we do with the drunken Sailor

Traditional

D m C

What shall we do with the drun-ken Sai-lor, what shall we do with the drun-ken Sai-lor,

D m C D m

what shall we do with the drun-ken Sai-lor, ear-ly in the mor-ning.

Gespielt wird auf einem nur zwei Oktaven umfassenden MIDI-Keyboard. Die d-Moll- und C-Dur-Skalen werden später kontrastierend farbig markiert. Auf einem Bildschirm sehen die Probanden die Audiospur ihres eigenen Klanges. An bestimmten Stellen ist diese Spur entsprechend der Keyboardmarkierung eingefärbt, sodass die Wechsel zwischen den beiden Harmonien mitverfolgt werden können und sich die jeweils aktuelle Skala für die harmonische Improvisation erschließt. Die Spielweise wird durch eine Videokamera dokumentiert, die auf Keyboard und Hände gerichtet ist.

Versuchsplan (Dauer: ca. 15 Minuten)

Impro 1 Freie Improvisation zum Playback (= Ausgangspunkt)

Musiktheoretische Intervention: Erläuterung der harmonischen Struktur

Impro 2a Gebundene Improvisation mit Dreiklangsgerüst (nur 1 Dreiklang)

Impro 2b Gebundene Improvisation mit Dreiklangsgerüst (2 Dreiklänge im Wechsel)

Impro 3 Gebundene Improvisation mit Orientierung am Dreiklangsgerüst, „Zwischentöne“ erlaubt

Impro 4 Versuch einer Synthese: von der (erneut) freien zur gebundenen Improvisation

Alle Aufgabenstellungen und Hinweise sind auf der 15-minütigen Tonspur vorproduziert. Einzig während der theoretischen Intervention wird die vorproduzierte Anfangsphase *What shall we do with the drunken sailor* mit dem Live-Spiel der entsprechenden Keyboard-Tasten unterstützt.

Detailliertes Protokoll der sog. „musiktheoretischen Intervention“

Stimme RL: *Das Lied beginnt so: (gesungen) „What shall we do with the drunken sailor.“*

dazu Keyboard: Anleitende/r spielt die Melodie bis „sailor“ mit. Am Ende werden alle drei verwendeten Töne (=d-Moll-Dreiklang) gehalten.

Stimme RL: *Was ergeben diese drei Töne? Wenn du es weißt, sag' es. Deine Antwort wird aufgenommen. (4 Sek. Pause.) Antwort: einen Dreiklang, also einen aus drei Tönen bestehenden Zusammenklang. Wie du siehst, ist zwischen den drei Tönen immer eine Taste ausgelassen (der/die Anleitende deutet mit freier Hand auf die Intervalle). So sieht ein typischer Dreiklang auf dem Klavier aus.*

Wir markieren die Dreiklangstöne auf den Tasten grün. (Es folgt Anweisung für Impro 2a.)

Interpretation des Versuchs

Von zentraler Bedeutung ist der Vergleich des Spiels vor und nach der Intervention des Anleitenden (Information über die Dreiklangsstruktur im Song) mit Blick auf drei mögliche Konsequenzen:

- A Ist die musiktheoretisch bedingte Restriktion nachhaltig (d. h. nach der Intervention verliert der Schüler die Fähigkeit, so frei wie zuvor zu improvisieren)?
- B Ist kurzfristig eine Synthese aus vorhandenen Fähigkeiten und dem soeben erworbenen theoretischen Wissen möglich?
- C Erweist sich das strukturelle Wissen gar als Impuls (d.h. tauchen ganz neue Spielweisen auf)?

Es handelt sich hier um den Versuch einer Förderung kreativer Leistung bei Kindern, die noch kein musikalisch-strukturelles Reflexionsvermögen entwickelt haben. A.C. Lehmann sieht Förderungen dieser Art in der Nähe einer „Hinführung auf ‚erwachsene‘ Muster von Kreativität hin“ (Lehmann 2005, 945). Das hieße in Bezug auf die oben spekulierten Konsequenzen:

- A würde bestätigen, dass eine solche Hinführung stattgefunden hat – negativ konnotiert.
- C würde dies ebenfalls bestätigen – jedoch positiv konnotiert,
- B würde dies widerlegen oder zumindest für eine Differenzierung sprechen.

Vorliegende Daten

- MIDI-File des Keyboardspiels
- Video mit Perspektive auf Keyboard/Hände
- Fragebögen, die vor und während der Einzelversuche ausgefüllt wurden

Vorrangig ist die kategoriale Auswertung der MIDI-Files. Multimediales Hauptwerkzeug ist dabei das Fenster *Matrix-Editor* (auch „*Pianorolle*“) im Studioprogramm *Logic* mit eingeblendetem synchronisierten Videofenster. Da jeder Schüler von Beginn an eine eigene Spielweise ausprägt, werden die Daten erst sukzessive aufeinander bezogen.

3 Ergebnisse

3.1 *Zwei Datenkomplexe*

Die Frage nach der Effizienz musiktheoretischer Lehre verbindet sich hier mit kreativen Tätigkeiten, d.h. mit Vorgängen, die auf individuelle Voraussetzungen der Probanden zurückgehen. Aus pragmatischen Gründen bezeichnen wir die Spielweise, wie sie sich bei der anfänglichen freien Improvisation (Impro 1) offenbart, dem Schüler/der Schülern gegenüber als dessen/deren „Stil“. Alle Kriterien, die im Datenkomplex „Stil“ aufgeführt sind, werden in den übrigen Improvisationsrunden jeweils erneut überprüft.

Datenkomplex A: „Stil“

- Finger (einer oder mehrere Spielfinger pro Hand)
- Hände (eine/beide; div. Unterkategorien zur Kombination der Hände sowie zu deren Abstand)
- Anschlagsdauer
- Anschlagsstärke (gemäß MIDI-Skala von 1-128)
- Anzahl der Anschläge
- Verwendete Tasten; verwendete schwarze Tasten
- Mehrstimmiges Spiel (einhändig/beidhändig)
- Struktur, vom Song unabhängig: Puls, Melodik, Phrasierung
- Struktur in Beziehung zum Song: Metrik/Rhythmus, Melodik, Harmonik

Zweitens wird das Verhalten in Bezug auf den Lerngegenstand „Dreiklang“ erhoben, d. h. zum einen der Lernzuwachs korrekten Skalenspiels, zum anderen aber auch die Nutzung der ab Impro 3 gestatteten Freiheiten. Letztere betreffen Art und Umfang der hinzukommenden „Zwischentöne“ (ab Impro 3) und die Dominanz entweder des freien oder des gebundenen Improvisierens (in Impro 4).

Datenkomplex B: Dreiklangspiel

- Mündl. Auskunft, ob das gezeigte Gebilde (Dreiklang) und der Fachbegriff schon bekannt sind
- Impro 2b: Verhältnis korrekter und fehlerhafter Töne
- Impro 3 u. 4: Verhältnis von Dreiklangstönen und akkordfremden Tönen („Zwischentönen“); Arten von „Zwischentönen“
- Impro 4: Zeitpunkt des Endes der freien Improvisation (= Einbeziehung des Dreiklangsgerüsts)
- Abschließendes Urteil (Fragebogen), ob das Beachten von Dreiklangsgerüsten *hinderlich* oder *hilfreich* ist (Erläuterung s. u.)

Auswirkung instrumentaler Vorbildung

In der betreffenden Schulklasse haben 4 von 26 Schüler/-innen eine mehrjährige instrumentale Vorbildung (2x Schlagzeug; 2x Klavier). Was die Bewältigung der Aufgaben zur dreiklangsgebundenen Improvisation anbelangt, unterscheidet sich diese Gruppe nicht nennenswert von den Mitschülern. Allerdings fällt bei der anfänglichen freien Improvisation (Impro 1) auf, dass...

- ... beide **Schlagzeuger** nicht nur die erwartungsgemäße metrische Sicherheit zeigen; ihrem Spiel sind auch vage Ansätze einer melodischen Motivik zu entnehmen. Möglicherweise sind eingeübte Abfolgen rechts- und linkshändiger Impulse dafür verantwortlich, dass auch bei beliebigen Handpositionen hier wiederholbare diastematische Strukturen entstehen.
- ... beide **Tasten-Instrumentalistinnen** zuerst Fragmente offenbar vertrauter Etüden (ohne Zusammenhang mit dem Song *Drunken sailor*) spielen, bevor sie sich dem freieren Spiel widmen.

3.2 Musiktheoretische Intervention und nachfolgendes Improvisieren mit Dreiklangsgerüsten (Impro 2a/2b und 3)

Während das Spiel in einem einzigen markierten Dreiklangsgerüst (Impro 2a: d-Moll über zwei Oktaven) von allen Schülern fast fehlerfrei umgesetzt wird, ist der Wechsel zwischen zwei Dreiklängen zum richtigen Zeitpunkt deutlich anspruchsvoller (Impro 2b: d-Moll/c-Dur). Auch er wird aber tendenziell mit hoher Trefferquote bewältigt (s. Abb. 1).

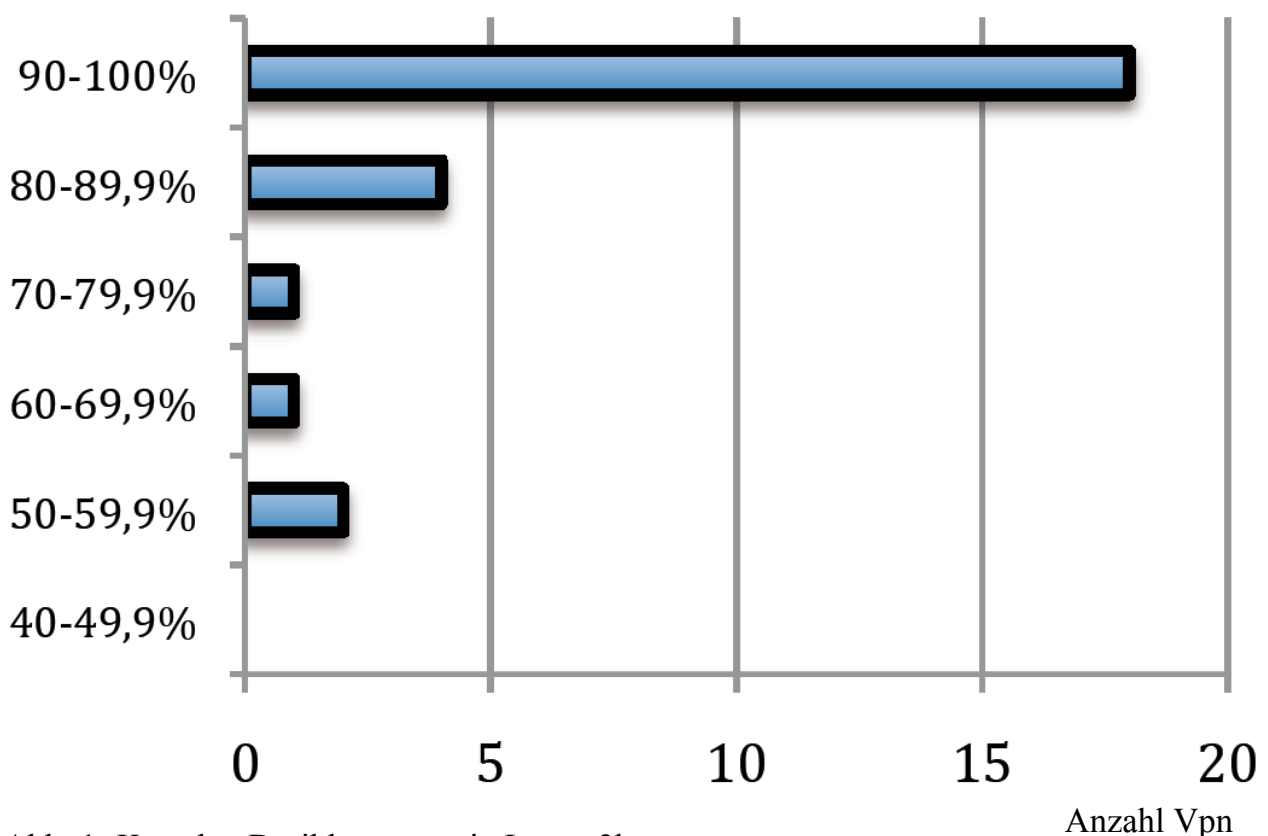


Abb. 1: Korrekte Dreiklangstöne in Impro 2b

Die Analyse der persönlichen Spielweise gemäß Datenkomplex A („Stil“) zeigt, dass trotz des massiven Eingriffs durch musiktheoretische Erläuterung und veränderte Aufgabenstellung das Verhalten der meisten Schüler/-innen in Bezug auf viele der hier gewählten Kategorien in Verlauf der Improvisationen 1 bis 3 ähnlich bleibt (s. Abb. 2). Berechnungsgrundlage der Entscheidung, ob die persönliche Spielweise als „stabil“ bezeichnet wird, ist, dass die Werte in mindestens 2/3 der Kategorien ohne signifikante Veränderungen bleiben.

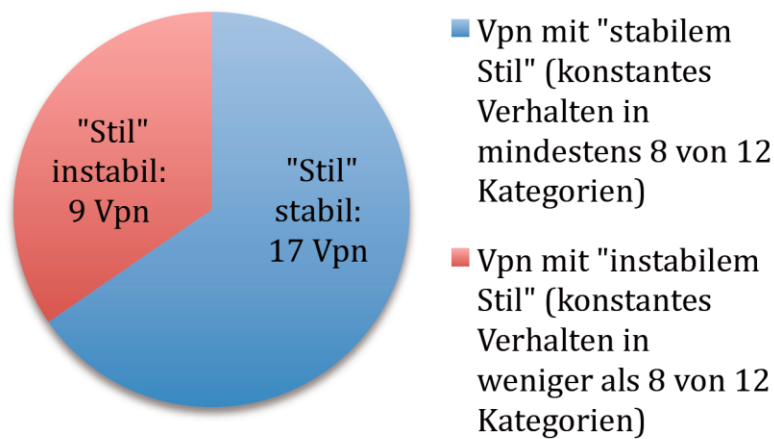


Abb. 2: Entwicklung der persönlichen Spielweise („Stil“) in Impro 1 bis 3

Allerdings kann die Klassifizierung als „stabil“ bei identischen Summen von sehr unterschiedlicher Qualität sein, weil im Datenkomplex A („Stil“) qualitativ sehr heterogene Kriterien versammelt sind, die – unabhängig von den Versuchspersonen – eine unterschiedliche Sensibilität für Veränderungen aufweisen. So liegt es beispielsweise nahe, dass eine mit den beiden Zeigefingern spielende Versuchsperson bei dieser Technik bleibt, während sie sich aber variabel in der Beziehung zur Songmelodie (Mitspielen von Melodietönen) entwickelt.

Die häufigsten Veränderungen im Verlauf der Improvisationen 1 bis 3 sind in folgenden Kategorien zu beobachten:

- Anschlagsstärke (Schwankungen auf der MIDI-Skala ab 10) 16 von 26 Vpn
- Beziehung zur Melodik des Songs (ja/nein) 13 von 26 Vpn
- Zahl der Anschläge (Schwankungen ab 10 Anschlägen) 13 von 26 Vpn

In diesen Kategorien verändert sich das Spiel hingegen weniger:

- Anschlagsdauer (tendenziell unter/über 1/8-Notenlänge) 6 von 26 Vpn
- Mehrstimmiges Spiel (ja/nein) 6 von 26 Vpn
- Metrische Struktur (erkennbar: ja/nein) 4 von 26 Vpn

Bei einem Teil der Kategorien lässt sich zur festgestellten Veränderung eine Bewertung vornehmen (als entweder positive Entwicklung oder aber Einschränkung durch musiktheoretisches Input). Interpretiert man

- das Hinzunehmen mehrerer Spielfinger oder der zweiten Spielhand,
- eine zunehmende mittlere Anschlagsstärke (sofern kein offensichtlich beabsichtigtes *piano*),
- eine Zunahme des verwendeten Tonvorrats,
- das Hinzukommen mehrstimmiger Spieltechniken,
- die Stabilisierung der Metrik,
- das Mitspielen von Tönen und rhythmischen Motiven aus dem Song

als Kennzeichen einer positiven Entwicklung, so lässt sich insgesamt eine erfreuliche Bilanz ziehen. Denn für die Mehrzahl der Kategorien mit „stilistischer Instabilität“ ist dann eine solche positive Entwicklung nachweisbar (beispielsweise in allen 13 Fällen bezüglich mitgespielter Songmelodik und in 12 der 16 Fälle bezüglich der Anschlagsstärke). Es lässt sich sogar sagen, dass keine einzige der 26 Vpn unter den instabil ausgeprägten Kategorien mehr Verringerungen als „Entwicklungen“ erkennen lässt.

Bestimmte Muster von Veränderungen, die wegen ihrer Individualität nur zum Teil in Datenkomplex A erfasst sind, kehren mehrfach wieder und bedürfen einer gesonderten Betrachtung.

	Spielweise in Impro 1		Spielweise in Impro 2 und 3
(a)	„Streuende“ Verteilung der Töne innerhalb eines bestimmten Ambitus	⇒	Verteilung der Töne in Wellenbewegungen
(b)	Hände konstant voneinander entfernt	wird zu	Überkreuzungen der Hände
(c)	Kaum erkennbare Metrik	⇒	Metrische Sicherheit
(d)	Legato-Tonleitern		Legato-Dreiklänge
(e)	Weiß und schwarze Tasten		Ausschließlich weiße Tasten

Hier gilt es in zwei Gruppen von Fällen zu differenzieren: In Fall (d) zeigt die Vp eine Spielweise, die unter den Bedingungen der strengen Dreiklang-

Improvisation unmöglich fortgesetzt werden kann. Auch das in Fall (a) beschriebene Spiel ist zwar theoretisch möglich, aber in Bezug auf Dreiklänge so anspruchsvoll, dass dessen Fortsetzung im Rahmen der hier agierenden Klientel illusorisch wäre.

Anders verhält es sich mit (b) und (e). Das Verlassen der jeweiligen Spielweise ist nicht zwingend. Doch in beiden Fällen spielen vermutlich soziale Muster der schulischen Lehre eine Rolle. Die Aufforderung des Lehrers zum Spiel im Dreiklangsgerüst suggeriert ein Verwenden der gesamten Skala und nicht nur der äußeren Töne (Fall b). Und die zunehmende Markierung von ausschließlich weißen Tasten führt ebenfalls die Konzentration auf einen bestimmten Tonvorrat, sodass selbst bei Erlaubnis von „Zwischentönen“ (Impro 3) die schwarzen Tasten aus dem Blick geraten.

Spielweise (c) und ihr plötzlicher Wandel schließlich sind besonders auffallend. Hierin verbirgt sich Potenzial für das Erlernen von Improvisationstechniken. Dass hier ein blitzartiger Lernzuwachs auf dem Gebiet der Metrik vorliegt, ist unwahrscheinlich. Nahe liegt, dass die klare und visuell unterstützte Reduktion des Tonvorrats nicht nur als zusätzliche Anforderung zu interpretieren ist, sondern auch als Entlastung: Die Last der beliebigen Tonwahl entfällt und schafft Kapazität für eine metrische Verankerung der Improvisation.

3.3 Musiktheoretische Intervention und spielerische Freiheit (Impro 3 und 4)

Improvisation 3 gestattet die Hinzunahme so genannter „Zwischentöne“, die nicht Teil des jeweils aktuellen Dreiklangs sind. Es werden in der Aufgabenstellung keine bestimmten Arten von Zwischentönen (z. B. *transitus* oder *Wechselnote*) festgelegt.

Die meisten Schüler/-innen nutzen diese Lizenz intensiv, d. h. die Quote korrekter Dreiklangstöne sinkt nach den hohen Quoten in Impro 2b (Abb. 1) auf Werte, die mehrheitlich zwischen 50 und 70 Prozent liegen (Abb. 3). Knüpft man die „intensiv genutzte Spielfreiheit“ an das Kriterium, dass die individuelle Quote korrekter Dreiklangstöne von Impro 2b zu Impro um mindestens 20%-Punkte sinkt, so trifft dies auf insgesamt 20 der 26 Vpn zu.

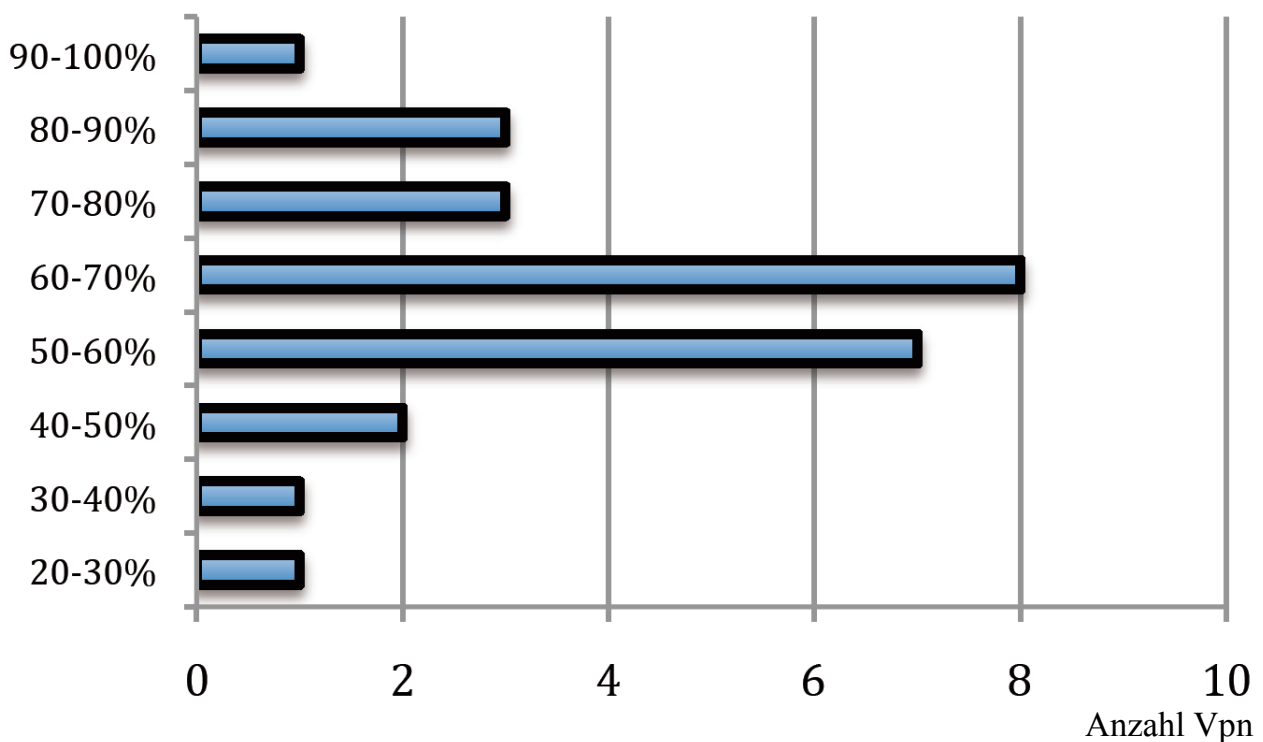


Abb. 3: Anteil korrekter Dreiklangstöne bei Improvisation mit „Zwischentönen“ (Impro 3)

Es werden sehr unterschiedliche Formen von „Zwischentönen“ gespielt. Die meisten resultieren aus dem Bemühen, trotz der Dreiklangorientierung lineare Melodiefragmente zu realisieren (belegt an der Summe der Versuche, in denen Durchgänge, Wechselnoten und Schritt/Sprung-Kombinationen erklingen, s. Abb. 4).

Verzerrt wird die Statistik in Bezug auf die Kategorie „Eingesprungene“, und zwar durch den Umstand, dass beim Verpassen eines Akkordwechsels automatisch Töne gespielt werden, die dem aktuellen Dreiklang nicht entsprechen. Sprunghaftes Dreiklangspiel zieht dann (wahrscheinlich oft ungewollt) Zwischentöne nach sich, die in die Kategorie „Eingesprungene“ fallen.

In wenigen Versuchen ist zu beobachten, dass der Aufgabenstellung Rechnung getragen wird, indem die Vp ihren Spielfluss unterbricht, nur um (sozusagen dem Lehrer zuliebe) einen „Zwischenton“ in einem völlig anderen Bereich der Tastatur zu setzen. Nach dem isolierten Ton kommt es dann zur Fortsetzung der ursprünglich begonnenen Melodie.

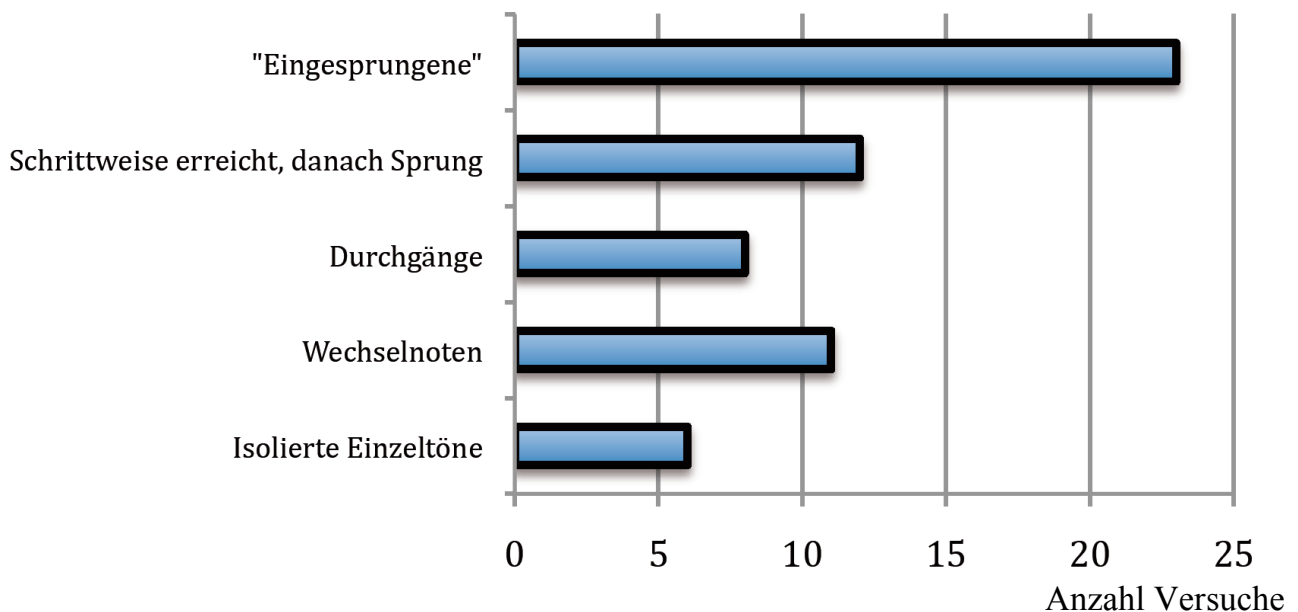


Abb. 4: Arten und Häufigkeit von „Zwischentönen“ in Impro 3

Improvisation 4 ist eine Aufgabe in mehreren Phasen. Als Vorbereitung darf die Versuchsperson die Aufnahme ihres anfänglichen freien Spiels (Impro 1) anhören. Impro 4 sieht dann vor, dass zuerst frei improvisiert wird und erst später – zu einem Zeitpunkt eigener Wahl – allmählich die Berücksichtigung von Dreiklangstönen erfolgen soll, bis schließlich eine Improvisation mit Dreiklangsgerüst erreicht ist. Auch hier sind „Zwischentöne“ erlaubt. Ziel der komplexen Aufgabe ist es herauszufinden, inwieweit ein vergleichbar freies Spiel wie am Anfang nach dem theoretischen Input (und bei markierten Tasten) überhaupt noch möglich ist und wie schnell die Probanden dazu übergehen, wieder Dreiklangstöne zu verwenden.

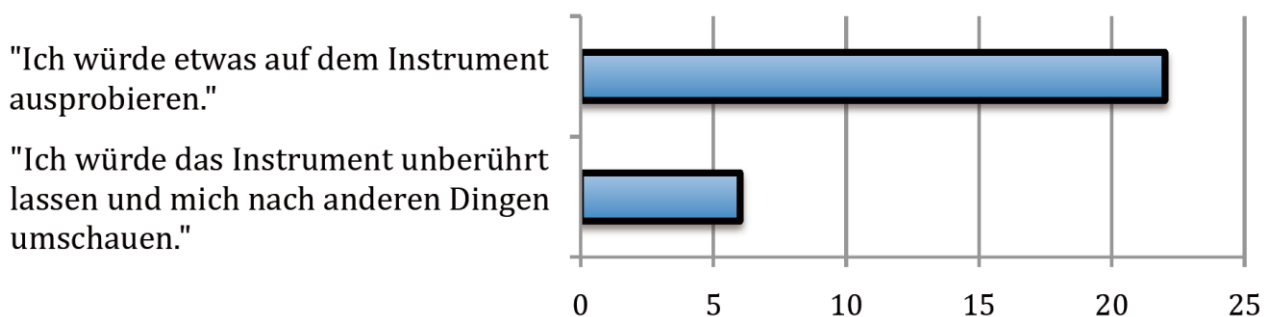
Trotz des Anhörens der eigenen freien Improvisation gelingt es in Impro 4 nur wenigen Schülern, ein ähnlich freies Spiel zu wiederholen. Obwohl sie wissen, dass das Song-Playback diesmal nicht auf 8 Takte begrenzt ist, sondern unbegrenzt wiederholt wird, ist eine Orientierung am Dreiklangsgerüst entweder sofort (6 Vpn) oder nach den ersten Takten vorhanden (2 Vpn nach 4 Takten; 3 Vpn nach 5 Takten, 2 Vpn nach 8 Takten). Diese von allen Kriterien losgelöste Art des Improvisierens wird ganz offenbar vermieden. Hier spiegeln sich die harten Selbsturteile der Schüler/-innen unmittelbar nach der ersten freien Improvisation (Impro 1) wider. Auf die Frage, wie sie ihr Spiel als Solo in einer Klassenmusizier-Situation bewerten, geben sich mehrere Schüler die Note 6 (Gesamtdurchschnitt: 3,9). Ob dies eher auf Hemmungen vor dem Präsentieren persönlichen Ausdrucks zurückzuführen ist oder auf die

ungewohnte Lernsituation ohne klare Leistungsbemessung, ist Spekulation. Nahe liegt eine Querverbindung zur o. g. Feststellung, dass sich mit der Einführung von Dreiklangstrukturen häufig eine Stabilisierung des gespielten Metrums verbindet. Insofern kann vermutet werden, dass das Spielen in definierten Rastern und die damit einhergehende Spielsicherheit von den Vpn positiv bewertet und bei freier Wahl bevorzugt wird.

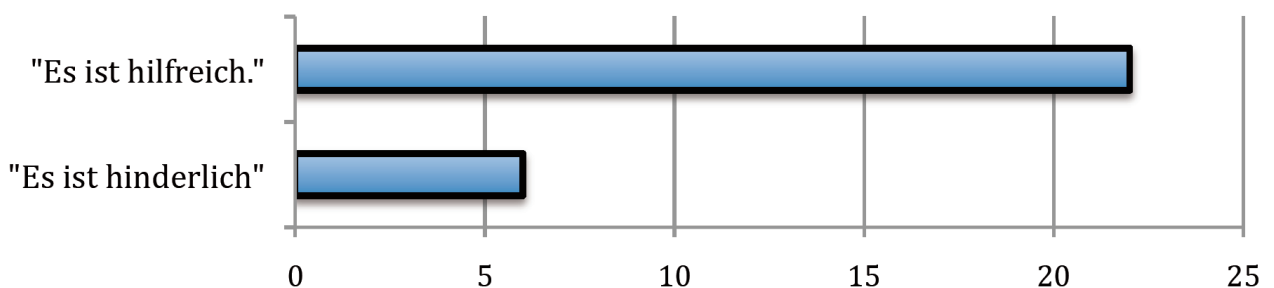
3.4 Selbsteinschätzung der Schüler/-innen und objektive Werte

Vor Beginn bzw. nach Abschluss des Experiments werden zwei Fragen gestellt:

1. „Stelle dir vor, du wärest zu Hause oder bei Bekannten gerade allein und würdest dich langweilen, und zwar in einem Raum, in dem außer vielen anderen Dingen auch ein Musikinstrument herumsteht (z. B. ein aufgeklapptes Klavier oder eine angelehnte Gitarre). Du wüsstest, dass das Spielen auf dem Instrument nicht verboten wäre. Wie würdest du dich verhalten?“



2. „Findest du das Beachten der Dreiklangs-Gerüsttöne eher hinderlich oder hilfreich, um Spaß am Keyboardspielen zu haben?“



Die Fragen beziehen sich auf unterschiedliche Dinge: zum einen auf die Neugier und innere Bereitschaft, unvorbereitet Musik zu machen, zum anderen

auf die Bewertung einer soeben kennen gelernten speziellen Art des Improvisierens. Gemeinsamer Aspekt ist eine gewisse Offenheit für das Annehmen musikalischer Angebote. Die Mehrheit der Schüler/-innen beantwortet die Fragen also so, wie es sich Lehrende wünschen dürften: Als experimentierfreudige und in Bezug auf die Lernmethode ‚dankbare‘ Versuchspersonen. Indes stellt sich bei näherem Hinsehen heraus, dass die beiden Minderheiten keineswegs identisch sind (nur 1 von 6 Vpn ist identisch). Es sind demnach zuvörderst experimentell eingestellten Schüler/-innen, die ihre Improvisationsfreude durch theoretische Aufgabenstellungen getrübt sehen.

Auch lässt sich nicht einfach von den Antworten auf das hier dokumentierte Improvisationsspiel schließen. Die Daten von nur 1 der 5 Vpn, die das Beachten der Gerüsttöne als „hinderlich“ einstufen, weisen (gemäß Datenkomplex A) insgesamt eine „stilistische“ Instabilität auf. Zudem liegt hier (wie bei den übrigen der 26 Vpn, s.o.) kein Fall vor, in dem unter den Instabilitäten die Verringerungen gegenüber den „Entwicklungen“ dominieren. Somit dürfen die Antworten auf die zweite Frage hier zunächst eher im rein hedonistischen Sinne als im Kontext vollzogenen Lernfortschritts gedeutet werden.

4 Fazit. Effizienz von Musiktheorie

Die Studie versucht die Effizienz musiktheoretischen Inputs (in Bezug auf ein bestimmtes Teilgebiet der Musiktheorie und gemäß einer ausgewählten Lernmethode) hinsichtlich der Entwicklung von Improvisationsmustern bei ca. 13-jährigen Schüler/-innen zu bestimmen. In den vorliegenden Ergebnissen spiegeln sich zwei Arten von „Effizienz“ wider.

4.1 „Effizienz“ – bezogen auf traditionelle Auffassungen von Lernzuwachs: „Die richtigen Töne spielen“

Ein Lernzuwachs, der darin besteht, dass ausschließlich die markierten Tasten in einem bestimmten Zeitfenster korrekt gespielt werden können, ist belegt (s. o. Abschnitt III-2). Mit ihm verbindet sich die grundsätzliche Erweiterung des musikalischen Erfahrungsraums insofern, als außer dem völlig freien Spiel erstmals auch ein Spiel mit Bezug zur Harmonik des Songs *What shall we do with the drunken sailor* erlebt wird. Freilich handelt es sich um eine weitgehend fremdgesteuert erworbene Erweiterung, die nicht weit über das vergleichbare „Malen nach Zahlen“ hinausgeht. Dennoch ist festzuhalten, dass nach Datenkomplex A keinerlei Indizien für eine Entmutigung durch den re-

striktiven Eingriff vorliegen. Das Spiel in schlichten Dreiklangsrastern wird motiviert ausgeführt und als Aufgabe bewältigt.

4.2 „Effizienz“ – bezogen auf eine Vernetzung von eigenem Background und Lerngegenstand: Individuelles Gestalten der Musik

Um Missverständnisse zu vermeiden: Das Improvisieren im Rahmen der Harmonielehre ist in Kompositionslehren seit dem 18. Jahrhundert üblich und darf also nicht das Attribut „innovativ“ für sich verbuchen. Nur scheint es bislang an Methoden zu mangeln, wie Improvisation auch Eingang in die basale Harmonielehre des regulären Schulunterrichts mit musikpraktischen Anfängern findet. Die hier vorgeschlagene Methode ist mehrschrittig und führt von einem freien Spiel über das strenge Dreiklangsraster zur erneuten Erweiterung des Tonvorrats. Ist sie effizient im Hinblick auf das Erlernen individuellen Gestaltens von Musik?

JA, was die Bereitschaft der Schüler/-innen angeht, über das erlernte strenge Dreiklangsgerüst hinaus auch wieder andere Töne zu spielen und (bei einem Teil der Probanden) Gerüst- und „Zwischentöne“ miteinander zu verbinden. Die in der Kürze der Zeit sicherlich teils zufällig entstandenen Zwischentöne könnten in einem nächsten Schritt beschrieben und klassifiziert werden, bevor sie gemäß persönlicher Vorlieben zunehmend beabsichtigt erklingen.

JA, was die unbewusste Beibehaltung von Elementen der individuellen Spielweise betrifft: ca. 2/3 der Probanden haben ihren anfänglichen „Stil“ insgesamt bewahrt; Veränderungen lassen sich mehrheitlich als Entwicklung deuten. Die Analysen sprechen dafür, dass sich hier schon kurzfristig eine enge Beziehung zwischen persönlicher Vorerfahrung und dem Lerngegenstand „Dreiklang“ abzeichnet.

NEIN, was die Reflexion des eigenen „Stils“ anbelangt. Das erneute freie Spiel in Impro 4 wird vermieden und – nicht überraschend – gelingt es kaum einem der Probanden, nach dem Hören der Aufnahme von Impro 1 Spezifika des eigenen freien Improvisierens zu erkennen und erneut in die Praxis umzusetzen.

Hier müssten weitere Phasen eingefügt werden, in denen verschiedene Aspekte des individuellen Spiels durch Mitschüler beschrieben, praktisch imitiert und variiert werden. Das gezielte Erhalten und Weiterentwickeln des eigenen „Stils“ wäre dann Prämisse sowohl bei der

Vertiefung von Dreiklangstrukturen als auch bei weiteren Phasen des ganz freien Improvisierens.

Abschließend sei noch einmal auf die Überlegungen zu „kindlichen“ und „erwachsenen“ Mustern von Kreativität rekurriert (These A. C. Lehmann). Nimmt man an, dass bei instrumental unerfahrenen und musiktheoretisch noch praktisch ahnungslosen Kindern auch noch im Alter von 13 Jahren dem Improvisieren kindlich-naive Aktionsmuster zugrunde liegen, so sind solche gemäß der präsentierten Studie nicht wiederholbar, nachdem die harmonische Struktur des Songs eingeführt und praktisch nachvollzogen worden ist. Es findet eine irreversible Hinführung zu erwachsenen Mustern von Kreativität statt (eine Hinführung im negativen Sinne, oben als Fall „A“ bezeichnet).

Gleichwohl kommt es in vielen Fällen individuell zu einer Weiterentwicklung der persönlichen Spielweise (also auch eine Hinführung im positiven Sinne, oben als Fall „C“ bezeichnet), aber eben nicht im Zuge von Naivität, denn eine solche wird im Moment des Erläuterns kompositorischer Strukturen unmöglich. Erinnert sei nur an das bewusste Setzen von akkordeigenen und akkordfremden Tönen in Impro 3.

Bestimmte Aspekte der Spieltechnik (jenseits der Melodiebildung) bleiben aber offenbar soweit unreflektiert, dass sie sich erhalten. Dies darf sicherlich auch „Synthese aus vorhandenen Fähigkeiten und dem soeben erworbenen Wissen“ (s. o. „Fall B“) genannt werden, ist jedoch letztlich nicht die eigentlich fokussierte Schnittstelle, wenn es um Improvisation und Harmonielehre geht.

Eine Folgestudie müsste sich wohl zum Level „erwachseneren“ Improvisierens bekennen und von Anfang an melodische Neigungen der Vpn stärker in den Blick nehmen. Komplexe Teile wie Impro 4 können dann entfallen und das Interpolieren von akkordfremden Tönen stärker fokussiert werden. Es steht außer Frage, dass für den Unterricht mit einer musikalisch wenig vorgebildeten Klientel noch Beiträge zur methodischen Schärfung benötigt werden, um die vielen wertvollen Elemente individuellen „Stils“ auch entlang eines musiktheoretischen Curriculums zu erhalten und weiterzuentwickeln. Mancherorts muss sich in der Lehre der Musiktheorie sicherlich erst noch die Erkenntnis verbreiten, dass bei Kindern ein wissensbasiertes musikalisches Verhalten eigentlich längst begonnen hat, wenn so genannter „Anfangsunterricht“ anberaumt wird. Improvisation kann helfen, dieses Wissen auszumachen und den schulischen Musikunterricht entsprechend zu optimieren.

Literatur

- Hans Bäßler und Ortwin Nimczik 2002: *Elementarlehre – ein elementares Missverständnis?*, in: Musik und Bildung, Heft 2 (April-Juni), S. 4-7.
- Marilyn G. Boltz 1991: *Some structural determinants of melody recall*, in: Memory and Cognition, Heft 19, S. 239-251.
- Herbert Bruhn 2005: *Wissen und Gedächtnis*, in: Thomas Stoffer und Rolf Oerter (Hg.), *Allgemeine Psychologie*, Göttingen, S. 537-590.
- Helga de la Motte-Haber 2005: *Musiktheorie und Psychologie*, in: Thomas Stoffer u. Rolf Oerter (Hg.), *Allgemeine Musikpsychologie*, Göttingen usw., S. 71-94.
- Folker Froebe 2006: „Musiktheorie und Vermittlung“ – Bericht über den VI. Kongreß der GMTH vom 6. bis 8. Oktober 2006 in Weimar (<http://www.gmth.de/zeitschrift/artikel/0114/0114.html>)
- Heiner Gembris 2008: *Musikalische Entwicklung im Erwachsenenalter*, in: Herbert Bruhn et al. (Hg.), *Musikpsychologie. Das neue Handbuch*, Reinbek bei Hamburg, S. 162-189.
- Wilfried Gruhn 2003: *Lernziel Musik*, Hildesheim.
- Christoph Hempel 2006: *Musik in Theorie und Praxis*, in: Peter Wicke (Hg.), *Duden Musik. Lehrbuch Gymnasium*, Berlin und Frankfurt/M., S. 42-109.
- Clemens Kühn 2002: *Musiktheorie zwischen Hochschule und Schule*, in: Diskussion Musikpädagogik, Heft 13, S. 7-21.
- Robert Lang 2008: *Musiktheorie im Unterricht. Impulse für eine umstrittene Fachdisziplin*, in: Frauke Heß und Jürgen Terhag (Hg.), *Musikunterricht Heute*, Bd. 7, Oldershausen, S. 70-84.
- Andreas C. Lehmann 2005: *Komposition und Improvisation. Generative musikalische Performanz*, in: Thomas Stoffer und Rolf Oerter (Hg.), *Allgemeine Psychologie*, Göttingen, S. 913-954.
- Marc Leman 2000: *An auditory model of the role of short-term memory in probe-tone ratings*, in: *Music perception* 17/2000, S. 481 ff.
- Ulrich Mazurkiewicz 2005: *Gegenstände des Musiklernens und Methoden des Musiklehrens*, Fernwald.
- Günther Noll 1973: *Zur Didaktik der musikalischen Improvisation (in der allgemeinbildenden Schule)*, in: Wolfgang Stumme (Hg.), *Über Improvisation*, Mainz u. a., S. 72-80.

Christoph Richter 2007: *Der Musikunterricht gehört den Schülern. Eine Anleitung zum selbständigen Umgang mit Musik*, darin: Kapitel „Theorie und Praxis“, Altenmedingen, S. 124-129.

Marco Ringel 2009: *Musiktheorie spielerisch erarbeiten*, Paderborn.

Richard Vella 2003: *Sounds in Space, Sounds in Time. Projects in Listening, Improvising and Composing*, London.

Bernhard Waldenfels 2001: *Lebenswelt als Hörwelt*, in: Karl Heinrich Ehrenforth (Hrsg.), *Musik – unsere Welt als andere. Phänomenologie und Musikpädagogik im Gespräch*, Würzburg.

Der BMBF-Forschungsschwerpunkt zu „Jedem Kind ein Instrument“ in Nordrhein-Westfalen und Hamburg

Einleitung

Jens Knigge & Ulrike Kranefeld (Universität Bielefeld)

Der vom Bundesministerium für Bildung und Forschung (BMBF) initiierte Forschungsschwerpunkt zu „Jedem Kind ein Instrument“ in Nordrhein-Westfalen (Ruhrgebiet) und Hamburg hat seit Beginn des Jahres 2009 die Arbeit aufgenommen. Der für zunächst vier Jahre (2009-2013) konzipierte Forschungsschwerpunkt ist unabhängig von den beiden JeKi-Programmen und mit einem Fördervolumen von jährlich rund 1 Mio. Euro Teil des BMBF-Rahmenprogramms zur empirischen Bildungsforschung.

Die JEKI-Programme in NRW und Hamburg erfordern aufgrund der mit ihnen verbundenen hohen Investitionen bei Bund und Ländern Forschungsarbeiten darüber, inwieweit zentrale Projektziele erreicht werden. Das BMBF fördert daher im Kontext des Forschungsschwerpunkts sieben größere Forschungsvorhaben. Diese bestehen aus insgesamt 14 Einzelprojekten, die in größtenteils interdisziplinären und universitätsübergreifenden Verbünden zu-

1 Die folgenden Personen sind Mitglieder des Konsortiums und haben die einzelnen Kapitel dieses Beitrags verfasst (eine gesonderte Nennung erfolgt jeweils auch bei den entsprechenden Kapiteln): Martin Andermann, Erich Beckers, Stephan Bongard, Martin Bosen, Veronika Busch, Monika Cloppenburg, Dorte Engelmann, Timo Fischinger, Emily Frankenberg, Kamala Friedrich, Christine Groß, Silvia Höllmüller, Sophie Ibrom, Jens Knigge, Ulrike Kranefeld, Gunter Kreutz, Sabrina Kulin, Katharina Lehmann, Andreas Lehmann-Wermser, Susanne Naacke, Anne Niessen, Sonja Nonte, Manfred Nusseck, Julia Rheingans, Ingo Roden, Peter Schneider, Renate Schneider, Melanie Schönbrunn, Klaudia Schulte, Knut Schwippert, Annemarie Seither-Preisler, Stefan Seither, Julika Wagschal.

sammenarbeiten. Beteiligt sind hierbei über 25 Wissenschaftler aus unterschiedlichen Disziplinen wie der Musikpädagogik, der systematischen Musikwissenschaft, der Musikpsychologie, den Neurowissenschaften und der Erziehungswissenschaft. Um die Kooperation zwischen den Forschungsprojekten zu fördern und den Kontakt der Wissenschaftler untereinander und zu den beteiligten Akteuren aus Musikschulen, Grundschulen, Ministerien und den JeKi-Programmen zu unterstützen, wurde vom BMBF eine Koordinierungsstelle eingerichtet, die an der Universität Bielefeld angesiedelt ist. Um eine Überforschung einzelner Schulen zu vermeiden, werden die jährlichen Erhebungen von der Koordinierungsstelle in Zusammenarbeit mit dem zentralen Datenpool an der Universität Bremen koordiniert. Beide sind als Serviceeinrichtungen integraler Bestandteil des Gesamtprogramms. Jährliche Netzwerktagungen und ein regelmäßiges Angebot für die beteiligten Nachwuchswissenschaftler unter Einbeziehung externer Experten dienen dem intensiven inhaltlichen und methodischen Austausch der Wissenschaftler untereinander und der systematischen Nachwuchsförderung innerhalb des Forschungsschwerpunkts.

Die Forschungsprojekte sind über das wissenschaftliche Erkenntnisinteresse hinaus bestrebt, ihre Ergebnisse in die Praxis zurückzuführen und dadurch einen Beitrag zur kritischen Reflexion und zu möglichen Weiterentwicklungen der JeKi-Programme zu leisten.²

1 Forschungsverbund AMseL: Audio- und Neuroplastizität des musikalischen Lernens³

Anders als bei den meisten schulischen Leistungen besteht beim aktiven Musizieren ein unmittelbares Feedback zwischen Handlung und dem Erleben des eigenen Erfolges. Auf diese Weise entsteht eine natürliche Tendenz, sowohl die eigenen Fähigkeiten optimal einzusetzen, als auch deren Wahrnehmung zu verbessern. Der Trend zur Optimierung lässt starke Lerneffekte erwarten, welche sich auch in entsprechenden neuronalen Veränderungen niederschlagen sollten (Overy 2004; Zatorre 2005). Trotz des öffentlichen Interesses an der

2 Weitere Informationen zum Forschungsschwerpunkt und den einzelnen Projekten sind im Internet unter der Adresse www.jeki-forschungsprogramm.de abrufbar.

3 Internet: <http://www.am-sel.org>

Frage, wie sich aktives Musizieren auf Gehör und Gehirn auswirkt und zahlreicher neurowissenschaftlicher Untersuchungen an Musikern ist noch wenig über die Dynamik lernbedingter plastischer Veränderungen bekannt (Jäncke 2008; Spychiger 2008). Das Ziel dieser für einen Zeitraum von 3 Jahren konzipierten Studie ist es, in kombinierten Quer- und Längsschnittsvergleichen zu untersuchen, wie sich regelmäßiges aktives Musizieren im Rahmen des Schulunterrichts auf die Sensibilisierung des Gehörs, auf allgemeine kognitive Fähigkeiten (Intelligenz, Kreativität, Aufmerksamkeit, Lesen, Rechtschreiben, Mathematikleistungen) und auf neuronale Verarbeitungsprozesse auswirkt. Zur Untersuchung sollen sowohl psychoakustische und psychologische Tests als auch neurologische Messungen (Magnetencephalographie: MEG; Magnetresonanztomographie: MRT) eingesetzt werden. Insgesamt sollen 360 Kinder getestet werden, wobei die Hälfte an dem Programm *Jedem Kind ein Instrument (JeKi)* teilnimmt. Neben schulisch unauffälligen Kindern, die außerhalb des JeKi-Programms entweder gar nicht oder sehr viel musizieren, sollen insbesondere Kinder mit Lese-Rechtschreibschwäche (Legasthenie) und dem Aufmerksamkeits(hyperaktivitäts)defizitsyndrom (AD(H)S) ohne außerschulische musikalische Förderung untersucht werden. Die Rekrutierung der Kinder mit Legasthenie und AD(H)S stellt eine besondere Herausforderung dar. Die Prävalenz liegt in der heutigen Gesellschaft bei ca. 3%-8% (bei Jungen stärker als bei Mädchen) mit einer zunehmend steigenden Tendenz. Der positive Einfluss musikalischer Früherziehung erscheint uns besonders bei Legasthenikern wichtig zu sein, da etwa 50% aller betroffenen Kinder auditive Defizite aufweisen (Hämäläinen et. al. i. Druck). Wir erhoffen uns zudem von den Ergebnissen wichtige Hinweise darauf, ob frühe musikalische Förderung auf andere schulische Lernbereiche Einfluss nimmt.

A Teilprojekt: Elementare Hörleistungen und die Wahrnehmung von Instrumentalklängen

Peter Schneider, Renate Schneider, Martin Andermann, Dorte Engelmann, Christine Groß & Sophie Ibrom (Universität Heidelberg)

Die musikalische Tonhöhe von harmonisch komplexen Tönen unterscheidet sich um bis zu drei oder vier Oktaven, wenn derselbe Ton unterschiedlichen Hörern vorgespielt wird (von Helmholtz 1885; Schneider 2005; Seither-Preisler 2007, 2008). Darüber hinaus zeigten Instrumentalensembles, Musikschulklassen, Chöre, Fraktionen von Sinfonieorchestern oder Lehrer und Schüler, die gut zueinander passen, überraschend homogene Hör- und Klangwahrnehmungseigenschaften (Schneider 2009). Diese vorläufigen Ergebnisse

lassen sich sowohl anlagebedingt als auch lernbedingt deuten. Dabei könnten entweder die Wahl eines bestimmten Instrumentes, die individuelle Klangwahrnehmung von Instrumentalklängen (Reuter 1995), das Lehrer-Schüler-Verhältnis oder sogar die Eignungsprüfung zur Aufnahme an eine bestimmte Musikschule durch die Art des Hörens beeinflusst werden. Umgekehrt könnte sich die Art des Hörens aber auch aus der Praxis mit einem bestimmten Instrument oder einem speziellen musikpädagogischen Unterrichtsansatz ergeben (Jäncke 2008).

Im Heidelberger Teilprojekt sollen die Lernvorgänge auf elementarer Ebene nach folgenden Gesichtspunkten untersucht werden: (1) Welche Höreigenschaften bringen die Kinder bereits mit? (2) Welche Veränderungen lassen sich aufgrund des JeKi-Unterrichts im Gehirn nachweisen? (3) Lassen sich die entsprechenden Charakteristika auch mit Hilfe psychoakustischer Tests nachweisen? (4) Wie bilden sich besonders intensives Musizieren oder Hörstörungen neurologisch und psychoakustisch ab? Welche Auswirkungen hat das Lernen eines Instrumentes langfristig auf die individuelle Klangwahrnehmung und die damit einhergehende neuronale Informationsverarbeitung? (5) Prägt die Klangeigenschaft eines Instruments im Laufe der Zeit das kindliche Gehör oder gibt es Instrumente, die für bestimmte Kinder z.B. mit einer speziellen Wahrnehmungsstörung besonders gut geeignet sind? Unser Ziel ist es, systematisch zu untersuchen, ob musikalische Früherziehung aufgrund der JeKi-Intervention zu einer Sensibilisierung des Gehörs sowie zu neuroplastischen Veränderungen bestimmter Gehirnstrukturen und -funktionen führt und eine Verbesserung allgemein-kognitiver Leistungen mit sich bringt. Außerdem soll untersucht werden, ob Kinder aufgrund der JeKi-Intervention im Vergleich zu Kontrollkindern ein unterschiedliches Entwicklungsmuster im musikalischen Lernen zeigen und ob mögliche Leistungsvorteile der JeKi-Kinder durch intensive außerschulische musikalische Förderung weiter verstärkt werden.

B Teilprojekt: Komplexe Hörleistungen im Zusammenhang mit Sprache und Musik

Annemarie Seither-Preisler, Stefan Seither & Silvia Höllmüller (Universität Graz)

Im Mittelpunkt des Grazer Teilprojekts steht die Untersuchung möglicher durch JeKi bedingter Veränderungen höherer auditiver Wahrnehmungsfunktionen, die eine wesentliche Voraussetzung für das Sprach- und Musikverständnis darstellen. Dabei soll untersucht werden, wie die Fähigkeit des Erkennens

bedeutungstragender akustischer Merkmale (auditive Szenenanalyse) durch aktives Musizieren beeinflusst wird. Im Rahmen des Projekts werden neben bereits etablierten Verfahren auch selbst entwickelte Hörtests entsprechend den aktuellen Fragestellungen angepasst und erprobt, um auditive Ursachen von Legasthenie zu identifizieren und durch musikalisches Training hervorgerufene Veränderungen der Wahrnehmung zu messen.

Darüber hinaus werden mögliche Wirkungen der JeKi-Intervention auf höhere kognitive Funktionen (Intelligenz, Kreativität, Aufmerksamkeit) und auf spezielle Domänen (Lesen und Rechtschreiben, Sprachverstehen, mathematische Fähigkeiten, musikalisches Vorstellungsvermögen) untersucht (Altenmüller 2008; Jäncke 2008). Derartige Effekte werden auch auf neuroanatomischer und -physiologischer Ebene für spezielle Gehirnbereiche analysiert. Insbesondere soll geklärt werden, ob musikalisches Lernen einen direkten Einfluss auf Gehirnstrukturen und -funktionen hat, die wichtig für mathematische und kreative Fähigkeiten sind.

Von den Ergebnissen beider Teilprojekte sind umfassende Erkenntnisse zu erwarten, auf welche schulisch relevanten Bereiche JeKi langfristig Einfluss nimmt.

2 Forschungsverbund GeiGE: Gelingensbedingungen individueller Förderung im Rahmen von JeKi – eine empirische Untersuchung an Grundschulen in Essen

Martin Bonsen & Monika Cloppenburg (Universität Münster); Anne Niessen & Katharina Lehmann (Universität Köln); Ulrike Kranefeld & Melanie Schönbrunn (Universität Bielefeld)

Eine der zentralen Zielsetzungen der musikpädagogischen Maßnahme „Jedem Kind ein Instrument“ (JeKi) besteht darin, Grundschülerinnen und -schüler aus bildungsfernen Schichten und insbesondere Kinder mit Zuwanderungsgeschichte zu fördern. Das von den Universitäten Münster, Köln und Bielefeld durchgeführte interdisziplinäre Verbundvorhaben greift dieses Anliegen auf, indem es als Evaluationsstudie JeKi im Kontext der jeweils allgemeinen und fachlichen grundschulpädagogischen Konzeptionen individueller Förderung auf verschiedenen Ebenen in den Blick nimmt:

- In der **ersten Teilstudie** (Münster, Martin Bonsen & Monika Cloppenburg) werden mit Hilfe quantitativer und qualitativer Verfahren Voraussetzungen,

Prozesse und Effekte der individuellen Förderung an 20 JeKi-Grundschulen der Stadt Essen untersucht, die unter pädagogisch besonders herausfordernden Bedingungen arbeiten. Im Zentrum der Teilstudie steht die Frage, inwieweit es den Grundschulen gelingt, JeKi in ihr spezifisches Schulprofil und ein schulinternes Curriculum individueller Förderung zu integrieren.

- Im **zweiten Teilprojekt** (Köln, Anne Niessen & Katharina Lehmann) rückt die Sichtweise der Lehrenden in den Mittelpunkt. Es wird mit qualitativen Methoden erforscht, über welches Konstrukt der „individuellen Förderung“ die beteiligten Grund- und Musikschullehrenden verfügen und wie sich dieses Konstrukt in der Teilnahme an JeKi konkretisiert.
- Die Instrumentenwahl der Schülerinnen und Schüler am Ende des ersten JeKi-Jahres wird im **dritten Teilprojekt** (Bielefeld, Ulrike Kranefeld & Melanie Schönbrunn) als exemplarisches Moment gedeutet, in dem sich individuelle Förderung in der Vermittlung von Entscheidungskompetenz im Unterricht konkretisieren kann. Auf der Mikroebene des Unterrichts werden Methoden der Instrumentenvorstellung analysiert und zu den am Ende des Jahres entwickelten Instrumentenpräferenzen der Schülerinnen und Schüler in Beziehung gesetzt.

Zentraler Forschungsgegenstand ist der Begriff der ‚individuellen Förderung‘; folgende Definition wird der Begriffsverwendung im vorliegenden Kontext zugrunde gelegt: „Unter individueller Förderung werden alle Handlungen von Lehrerinnen und Lehrern und von Schülerinnen und Schülern verstanden, die mit der Intention erfolgen bzw. die Wirkung haben, das Lernen der einzelnen Schülerin / des einzelnen Schülers unter Berücksichtigung ihrer / seiner spezifischen Lernvoraussetzungen, -bedürfnisse, -wege, -ziele und -möglichkeiten zu unterstützen“ (Kunze 2008, S. 19).

In **Teilprojekt 1** wird eine Stichprobe von 20 Grundschulen in eher strukturschwachen Stadtteilen der Stadt Essen untersucht. Die Schulen der Stichprobe sind pädagogisch besonders herausgefordert durch viele Schülerinnen und Schüler mit Migrationshintergrund und aus benachteiligten sozioökonomischen Verhältnissen. Hieraus ergeben sich der Betrachtungsfokus des Teilprojekts 1 und die grundlegenden Forschungsfragen:

- Welche Strategien, Prozesse und Entwicklungen zur Verbesserung der individuellen Förderung lassen sich an Schulen, unter Berücksichtigung ihrer jeweils spezifischen Herausforderungen, identifizieren?

- Welche Rolle spielt an den befragten Schulen das Projekt „Jedem Kind ein Instrument“ im Rahmen der individuellen Förderung?

Den theoretischen Hintergrund des Teilprojekts fasst das Prozess-Modell zur Qualitätssicherung im Bildungswesen von Ditton (2009, S. 609) zusammen. Dieses geht von engen Verknüpfungen von Voraussetzungen, innerschulischen Prozessen und den auf Schülerebene realisierten Ergebnissen aus. Individuelle Förderung ist ein wichtiger Bestandteil dieser Prozesse, welcher im Rahmen von Schul- und Unterrichtsentwicklung auf der institutionellen Ebene der Einzelschule optimiert und gestaltet werden soll. Schulentwicklung wird hier verstanden als vom Kollegium gemeinsam verantwortete und entwickelte Gestalt von Schule (Esslinger 2002, S. 23). Die jeweils spezifischen Herausforderungen und Voraussetzungen von Schule haben großen Einfluss auf die Gestaltung der schulischen und unterrichtlichen Prozesse, vermutlich auch auf die Intensität der individuellen Förderung. Aktuelle Schulleistungsstudien (z. B. TIMSS, PISA, IGLU) haben den Einfluss der sozialen Herkunft auf den Kompetenzerwerb von Schülerinnen und Schülern wiederholt empirisch belegt und konnten solche Hintergrundmerkmale von Schülerinnen und Schülern identifizieren, die einen großen Einfluss sowohl auf die pädagogischen Prozesse in der Schule als auch auf das Lernen der Schülerinnen und Schüler haben, wie z. B. das Bildungsniveau der Eltern oder die im Elternhaus genutzte Sprache (vgl. Bonsen et al. 2008). Zur empirischen Betrachtung dieser Prozesse in den JeKi-Schulen kommen unterschiedliche Forschungsmethoden zum Einsatz: explorative Interviews mit Schulleitungen und JeKi-Lehrenden sowie Fragebögen für Lehrkräfte, JeKi-Lehrende und Eltern der JeKi-Kinder. Zusätzlich werden Dokumente (z. B. das in NRW obligatorische Schulprogramm und schulinterne Curricula) analysiert. Schließlich sollen Portfolios den Stand und die Entwicklung der Schulen dokumentieren.

In **Teilprojekt 2** werden mit Hilfe eines dezidiert qualitativen Zugriffs Lehrende in den Blick genommen, die an den Sample-Schulen des Teilprojekts 1 im Essener Norden arbeiten. Die Fragestellung des Teilprojekts 2 lautet:

- Welche Rolle spielt das Konzept der individuellen Förderung in den Individualkonzepten von Musikschul- und Grundschullehrenden des JeKi-Projekts?

Dazu werden qualitative Interviews mit Grundschul- und Musikschullehrenden durchgeführt: In Experteninterviews werden die Gesprächspartner zu ihrem beruflichen Handeln befragt, in biographischen Interviews werden die

jeweiligen berufsbiographischen Hintergründe erhoben. Den theoretischen Hintergrund für die Erhebung stellt die Theorie der Individualkonzepte von Musiklehrenden dar (Niessen 2006). Die Interviewstudie ist im Sinne der „Grounded Theory Methodologie“ angelegt, was Einfluss auf die Auswertungsstrategien wie auf die Auswahl der Interviewpartnerinnen und -partner nimmt: Je nach Stand der Auswertung und Theoriebildung werden Anzahl und Auswahl weiterer Interviewpartner bestimmt (Glaser, Strauss 1998). In jedem Fall werden die Individualkonzepte zweier Lehrendenteams im ersten Jahr des Förderzeitraums erhoben und nach ca. anderthalb Jahren im Hinblick auf Veränderungen erneut untersucht; darüber hinaus werden im Sinne des Theoretical Sampling weitere Lehrende befragt.

Im **Teilprojekt 3** werden aus dem Sample des Teilprojekts 1 sieben Klassen begleitet. Im Zentrum stehen Unterrichtsanalysen zur individuellen Förderung im Unterricht des 1. JeKi-Jahres. Den Hintergrund bilden Hinweise auf einen Zusammenhang zwischen der Art und Weise der Instrumentenvorstellung und der späteren Instrumentenwahl (Byo 1991). Deshalb lauten zentrale Fragestellungen des Teilprojekts 3:

- Welche unterschiedlichen Erfahrungsräume werden den Schülerinnen und Schülern geboten, um in (intensiven) Kontakt mit den Instrumenten zu treten?
- Welche besonderen Bedingungen entstehen dadurch, dass JeKi-Lehrerinnen oder JeKi-Lehrer Musikinstrumente vorstellen, die sie nicht selbst studiert haben?
- Wie nutzen die Schülerinnen und Schüler diese Angebote der Instrumentenvorstellung?
- Welche Spuren hinterlassen die Instrumentenvorstellungen bei den einzelnen Schülerinnen und Schüler unmittelbar oder nach längerer Zeit bzw. im Kontext der Instrumentenwahl?

Eingesetzt werden videobasierte Methoden interpretativer Unterrichtsforschung (Kranefeld 2008): Der JeKi-Unterricht wird in sieben 1. Klassen wiederholt videografiert. Im Fokus stehen dabei jeweils die JeKi-Stunden, in denen die Geige, die Block- und Querflöte und Blechblasinstrumente als Beispiele für Instrumente mit unterschiedlicher Tonerzeugung vorgestellt werden. Ergänzend werden einzelne Stunden im Sinne mehrperspektivischer Unterrichtsforschung genauer untersucht (Niessen & Lehmann-Wermser 2006), wobei u. a. bei der rückblickenden Befragung der Schüler Standbilder oder

Videosequenzen im Sinne eines stimulated recall verwendet werden (Irion 2008).

Das Verbundvorhaben ist als formative Evaluation angelegt, die entwicklungsbegleitend zur Optimierung des Lernangebots beitragen soll. Dazu werden Erhebungsdaten an die Schulen zurückgemeldet und mit den Lehrkräften gemeinsam Instrumente zur Schulprogrammarbeit, zur Selbstreflexion sowie zur Schülerbeobachtung und -einschätzung entwickelt.

3 Forschungsverbund MEKKA: MusikErziehung, Kindliche Kognition und Affekt

A Teilprojekt: Stresserleben und Stressbewältigung

Stephan Bongard, Emily Frankenberg & Kamala Friedrich (Universität Frankfurt)

Das Frankfurter Teilprojekt befasst sich mit den emotionalen Auswirkungen des Musikunterrichts. Es wird untersucht, inwiefern das Musizieren zu einer Verringerung emotionaler Vulnerabilität und zu einer Verbesserung der emotionalen Selbstregulation beitragen kann. Ferner soll geprüft werden, inwieweit das regelmäßige Instrumentalspiel selbst eine emotionale Belastung mit Blick auf weitere Anforderungen des Schulalltags darstellt.

Wirkungen des Instrumentalspiels auf Denken, Fühlen und (soziales) Handeln von Kindern wurden bereits mittels Längsschnittstudien (z. B. Bastian 2000; Weber, Spychiger & Patry 1993) untersucht. Ziel des geplanten Teilprojektes ist es, Auswirkungen des JeKi-Instrumentenunterrichts auf spezifische Aspekte der Emotionsregulation im Kindesalter zu untersuchen und mit Hilfe geeigneter psychometrischer und biologischer Marker neu zu bewerten. Dabei rücken, stärker als in bisherigen Forschungsansätzen, differenzielle Aspekte der kindlichen Entwicklung in den Vordergrund.

Die Fragestellung soll innerhalb eines einfachen Kontrollgruppen-Designs mit Messwiederholungen untersucht werden. 250 Kinder der JeKi Gruppe werden bzgl. der abhängigen Variablen mit 250 Kindern einer Vergleichsgruppe kontrastiert. Kinder der Vergleichsgruppe sind solche, die kein Instrument erlernen und es zu Untersuchungsbeginn auch nicht geplant ist, dass sie ein solches erlernen werden.

Als abhängige Variablen werden erfasst:

psychometrisch

- Subjektive Belastungen und der kindliche Umgang mit diesen (SSKJ; Lohaus, Eschenbeck, Kohlmann & Klein-Heßling 2006)
- Grundlegende emotionale und soziale Erfahrungen in der Schule (FEES 3-4; Rauer & Schuck 2003)
- Kindliche Akkulturationsstile (eigene Entwicklung in Anlehnung an Bongard 2007)
- Dissoziative Erfahrungen (eigene Entwicklung in Arbeit)

behavioral/physiologisch

- Cortisol-Speichelkonzentration unter Ruhebedingungen
- Cortisol Stressreaktionen unter Belastungsbedingungen
- Blutdruck und Herzrate vor und während einer emotionalen Belastungssituation (wird nur bei einer Teilstichprobe von 30 JeKi- und 30 Vergleichskindern erhoben)

klinisch (halbstandardisierte Gespräche mit den Eltern)

- problematische Verhaltensweisen (CBCL 4-8; Arbeitsgruppe CBCL 1998)
- ärztliche Diagnosen und ärztlich verordnete Medikation

B Teilprojekt: Aufmerksamkeit und Gedächtnisleistung

Gunter Kreutz & Ingo Roden (Universität Oldenburg)

Das Oldenburger Teilprojekt befasst sich mit den kognitiven Auswirkungen des Musikunterrichts. Es wird untersucht, inwiefern das Musizieren zu verbesserten kognitiven Leistungen beitragen kann und ob sich diese Verbesserungen über verschiedene kognitive Prozesse hinweg ähnlich oder unterschiedlich darstellen.

Frühere Querschnitt- und Längsschnittstudien (siehe die Zusammenfassung von Jäncke 2009) bieten eine gemischte Befundlage. Bislang bleibt unklar, inwieweit Unterricht auf Musikinstrumenten in bestimmten Bereichen des Denkens, insbesondere musikspezifische Verarbeitung, oder darüber hinaus in

eher übergeordneten Domänen, etwa Aufmerksamkeit und Gedächtnis, Veränderungen bewirken kann.

Die Fragestellung des Teilprojekts setzt diesen Diskurs fort und soll somit zur Deutung musikinduzierter Veränderungen kognitiver Leistungen einen empirischen Beitrag leisten.

Als abhängige Variablen werden erfasst:

psychometrisch

- Aufmerksamkeits-Belastungstest zur Erfassung der kurzfristigen Aufmerksamkeit und Konzentrationsleistung (d2; Brickenkamp 2002)
- Kognitiver Fähigkeitstest für 1.-3. Klasse (KFT 1-3; Heller & Geisler 1983), Subtests 3 (Schlussfolgerndes Denken) und 4 (Rechnerisches Denken)
- Test zur Informationsverarbeitungsgeschwindigkeit (ZVT; Oswald und Roth 1987)
- Arbeitsgedächtnistestbatterie für Kinder von 5 bis 12 Jahren (AGTB 5-12; Hasselhorn, Schumann-Hengsteler, Grube, König, Mähler, Schmid, Seitz-Stein & Zoelch in Vorb.)
- Verbaler Lern- und Merkfähigkeitstest (VLMT; Helmstaedter, Lendt & Lux 2001)
- Audiationsfähigkeit und musikalisches Potenzial bei Kindern (IMMA; Gordon 1984)
- Zahlen-Merkfähigkeiten-Test (Eigene Entwicklung in Arbeit)

Die Datenerhebungen sollen zeitgleich für beide Teilprojekte zu drei Messzeitpunkten (Okt/Dez. 2009; Sept./Nov. 2010; März/Mai 2011) erfolgen.

4 Forschungsverbund SIGrun: Studie zum Instrumentalunterricht in Grundschulen

Die „Studie zum Instrumentalunterricht in Grundschulen“ (SIGrun) der Universitäten Bremen und Hamburg hat zum Ziel, den JeKi-Unterricht in Hamburg und in Nordrhein-Westfalen (NRW) aus verschiedenen Blickwinkeln zu beleuchten. Zu diesem Zweck werden sowohl quantitative als auch qualitative

Forschungsmethoden eingesetzt, die miteinander verknüpft werden sollen. Der SIGrun-Forschungsverbund besteht aus insgesamt vier Teilprojekten, die einen je spezifisch inhaltlichen Fokus aufweisen: 1) mögliche Transfereffekte des Programms auf kognitive, motivationale bzw. auf der Klassenebene klimatische Variablen, 2) Kooperationsbedingungen und Auswirkungen der Zusammenarbeit zwischen Musikschulen und Musiklehrkräften an den teilnehmenden Grundschulen, 3) Förderung der kulturellen Teilhabe von Kindern, 4) Gründe für die musikalische Präferenzentwicklung von Kindern.

Aufgrund der unterschiedlichen Situationen in NRW und Hamburg sind für jedes der beiden Bundesländer unabhängige Stichproben ausgewählt worden. Um den sog. „schooling effect“ (Stern 2006) zu verhindern, werden die in NRW an JeKi teilnehmenden Schülerinnen und Schüler mit Kindern verglichen, die eine sportliche Förderung erhalten. In Hamburg werden die Kinder an JeKi-Schulen mit Schülerinnen und Schülern an Schulen verglichen, die im Rahmen eines von der Hochschule für Musik betreuten Schulversuchs schon länger einen Schwerpunkt auf musikalische Erziehung gelegt haben. Zum ersten Messzeitpunkt werden rund 400 Schülerinnen und Schüler je Gruppe befragt. Die Gesamtstichprobe zum ersten Messzeitpunkt umfasst somit rund 800 Schülerinnen und Schüler je Bundesland sowie deren Eltern, Klassen-, Musik- und Musikschullehrkräfte sowie Schulleitungen. Je Bundesland soll eine einheitliche Klassen-Jahrgangskohorte mit einer Befragung im Jahr über die Zeit der Grundschule verfolgt werden. Die erste Erhebung ist am Ende der ersten Klasse vorgesehen, weitere Messzeitpunkte und eine Weiterführung ist für den Antragszeitraum bis in die vierte Klassenstufe geplant. Das Untersuchungsdesign lässt eine Längsschnittmodellierung unter Kontrolle von Klassenzusammensetzung und unter Berücksichtigung von Unterrichtsmerkmalen anhand von Mehrebenen-Messwiederholungs-Modellen zu. Zusätzlich werden einige Kinder, sowie deren Eltern und Lehrkräfte in ausgewählten Schulen qualitativ befragt.

A Datenpool der Studie zum Instrumentalunterricht in Grundschulen

Klaudia Schulte & Andreas Lehmann-Wermser (Universität Bremen)

Die Struktur des Gesamtvorhabens der Studie zum Instrumentalunterricht in Grundschulen und insbesondere die intensive Vernetzung der Teilprojekte ließen es sinnvoll erscheinen, einen sog. „Datenpool“ einzurichten: Damit ist es möglich, dass alle Teilprojekte auf einen umfassenden, belastbaren und spezifischen Bestand an Daten zurückgreifen können. Der besondere Gewinn des

Vorhabens liegt darin, dass die Teilprojekte aus verschiedenen subdisziplinären Perspektiven und mit einander ergänzenden Detailfragestellungen an einer gemeinsamen Stichprobe arbeiten werden, wobei sie sich an einer geteilten leitenden Fragestellung orientieren. Hierdurch und durch die längsschnittliche Anlage des Vorhabens sollen kumulative, über die Einzelvorhaben hinausgehende Ergebnisse erzielt werden.

B Teilprojekt: Transfer

Knut Schwippert & Sonja Nonte (Universität Hamburg)

Seit Rauschers, Shaws und Kys Publikation zu „Music and Task Performing“ im Jahr 1993, verbreitete sich weltweit der Befund, dass das Hören von Musik Mozarts zu besseren Werten bezüglich des räumlich-visuellen Vorstellungsvermögens in Intelligenztests führt. Seitdem wurde eine große Anzahl an quasi-experimentellen Studien durchgeführt und publiziert, die sich der Überprüfung dieses sogenannten „Mozart-Effekts“ zuwenden. Ein Aspekt der Transferforschung beruft sich demnach auf die Annahme, dass Musikhören Einfluss auf unterschiedliche Bereiche der Intelligenz, insbesondere auf das räumlich-visuelle Vorstellungsvermögen hat. Schellenberg (2003) fasst die Ergebnisse dieser Studien zusammen und kritisiert die empirische Fundierung und Evidenz der Befunde. Er kommt zu dem Schluss, dass nicht die Mozart-Musik an sich zu besseren Testleistungen führt, sondern dass jegliche von der Versuchsperson als emotional anregend wahrgenommene Intervention – hierzu zählt z.B. auch das Hören eines Hörspiels – einen positiven Effekt auf die Testleistung hat. Aufgrund dieser Befunde schreiben Nantais und Schellenberg (1999) diesen Effekt somit nicht der Musik Mozarts zu, sondern gehen davon aus, dass das Hören von Musik oder Hörspielen Einfluss auf den Erregungszustand und die Stimmung (*arousal-mood-theory*) nehmen kann. Diese wiederum beeinflussen die räumlich-visuellen Fähigkeiten.

Der zweite Strang der Transferforschung im musikalischen Kontext befasst sich mit Effekten des aktiven Musizierens auf kognitive, soziale, emotionale sowie motivationale Bereiche. Die große Anzahl an quasi-experimentellen Studien hat auch hier bisher nicht zu einem empirisch abgesicherten Erkenntnisgewinn beigetragen. Costa-Giomi (2004) konnte anhand einer Längsschnittstudie belegen, dass sich die schulische Leistung von Kindern, die über einen Zeitraum von drei Jahren Klavierunterricht erhielten, gegenüber der Kontrollgruppe nicht verbesserte. Insgesamt konnte aber eine Verbesserung des allgemeinen Selbstwertgefühls bestätigt werden. Auch in Deutschland und

der Schweiz wurden zunehmend Projekte mit Instrumentalunterricht an allgemeinbildenden Schulen implementiert und evaluiert (Weber et al. 1993, Bastian 2002, Schulten 2009). Die Befunde der Studien unterscheiden sich und sind außerdem häufig nicht eindeutig interpretierbar.

Einer der Forschungsschwerpunkte im Kontext von SIGrun befasst sich unter anderem mit Transfereffekten von aktiver musikalischer Betätigung auf Merkmale, die im Zusammenhang mit schulischer Leistung stehen, wie dem allgemeinen und musikalischen Selbstkonzept, der Schulzufriedenheit und dem Klassenklima. Weiterhin werden das Unterstützungsverhalten der Eltern, die sozialen und ökonomischen Hintergrundmerkmale sowie die kognitiven Grundfähigkeiten kontrolliert.

Die Evaluationsstudie ist als Längsschnittstudie mit vier Messzeitpunkten geplant und begleitet Kinder in Nordrhein-Westfalen und in Hamburg ab der ersten Klasse in einem jährlichen Turnus. Die Kinder werden dabei mit standardisierten Testverfahren getestet und kindgerecht mit Hilfe eines standardisierten Fragebogens befragt. Zusätzlich werden die Eltern ebenfalls mit standardisierten Fragebögen befragt. Die Kontrollgruppe ohne musikalische Förderung setzt sich aus Kindern zusammen, die eine Schule mit einem sportlichen Förderkonzept besuchen.

Eine zentrale Forschungsfrage des Projekts befasst sich mit den Zusammenhängen zwischen alternativen Förderkonzepten (Musik und Sport) und individuellen, auf den schulischen Kontext bezogenen Merkmalen, wie z. B. mit dem musikalischen, sportlichen aber auch schulischen Selbstkonzept sowie mit der Schulzufriedenheit und dem Klassenklima. Mit ersten Ergebnissen zu längsschnittlichen Zusammenhängen kann frühestens nach dem dritten Messzeitpunkt, im Jahr 2011, gerechnet werden.

C Teilprojekt: Kooperation

Knut Schwippert & Sabrina Kulin (Universität Hamburg)

Ein wesentliches Merkmal einer sich verändernden Schul- und Unterrichtskultur spiegelt sich in einer zunehmenden Zusammenarbeit mit außerschulischen Kooperationspartnern wider. Nach Helms (2002) zeichnet sich Kooperation durch die gemeinsame Nutzung und Ergänzung von Ressourcen aus, die sich auf Handlungsmöglichkeiten und pädagogische Kompetenzen beziehen. Auf diese Synergieeffekte baut ein Grundgedanke des Projekts JeKi (Jedem Kind ein Instrument), bei dem Kooperationen zwischen Grund- und Musikschulen

gepflegt werden, indem von Musikschullehrkräften in Räumlichkeiten der Schule Instrumentalunterricht erteilt wird.

Bisherige Forschungen zu Schulk Kooperationen fokussieren insbesondere die Zusammenarbeit von Schule und Jugendhilfe, wenngleich auch Musikvereine und kommerzielle Musikschulen nachweislich zu Kooperationspartnern von Schulen gehören (vgl. Arnoldt 2007; Behr-Heintze & Lipski 2005; Meyer-Clemens 2007). Evaluationen zur Kooperation von Schule und Musikschule liegen jedoch kaum vor.

Das Teilprojekt ‚Kooperation‘ geht zum einen der Frage nach, wie Kooperationsqualität im JeKi-Kontext definiert und gemessen werden kann sowie welche Faktoren mit dieser im Zusammenhang stehen. Zum anderen wird der Einfluss der Kooperationsqualität auf verschiedene schulische Aspekte wie beispielsweise das Schul- und Klassenklima oder Lehrermerkmale erforscht. Hierzu gehört unter anderem das Erfassen des musikalischen Selbstkonzepts der Lehrkräfte. Auf Grund bisheriger Erkenntnisse wird erwartet, dass sich das musikalische Selbstkonzept der Grundschullehrkräfte durch die Kooperationen verändert. Dafür sprechen beispielsweise die verschiedenen Ansätze und Ziele von Musikunterricht der Grund- und Musikschule (vgl. Bastian 2001; Helms 2002) sowie Forschungsergebnisse, bei denen gezeigt wurde, dass sich Lehrkräfte durch den Einsatz von Schulsozialarbeitern in ihrer pädagogischen Freiheit eingeschränkt fühlten sowie Zweifel an der eigenen Befähigung bekamen (vgl. Behr-Heintze & Lipski 2005).

Um diesen Fragen nachzugehen, werden quantitativ mit Hilfe von standardisierten Fragebögen Merkmale erhoben, die auf Grund theoretischer Diskussionen und wissenschaftlicher Forschungen einen Zusammenhang mit der Kooperationsqualität erwarten lassen (vgl. Czerwanski 2003; Jerusalem & Schwarzer 1991). Da der Zusammenarbeit zwischen Kooperationspartnern außerdem bestimmte (Beziehungs-) Strukturen zu Grunde liegen, werden jene durch den Einsatz der sozialen Netzwerkanalyse innerhalb des Fragebogens erfasst. Dadurch können sowohl die Strukturen, in denen die Akteure im Kooperationskontext eingebettet sind, als auch die Beziehungen der Beteiligten untereinander untersucht werden (vgl. Jansen, 2006). Die quantitativ erhobenen Daten werden durch leitfadengestützte Interviews ergänzt.

D Teilprojekt: Kulturelle Teilhabe

Andreas Lehmann-Wermser & Susanne Naacke (Universität Bremen)

Das Teilprojekt ‚Kulturelle Teilhabe‘ wird sich mit Zugangswegen zu Kultur bzw. Musik im Rahmen der JeKi-Programme in Hamburg und im Ruhrgebiet beschäftigen. Kulturelle Teilhabe wird dabei als Voraussetzung von kultureller Bildung und somit als Bestandteil von Allgemeinbildung verstanden (vgl. Fuchs & Bockhorst 2008). Die Befunde der PISA-Studien verdeutlichen, dass es dem deutschen Bildungssystem nach wie vor nicht in vollem Umfang gelingt, die durch ungünstige familiäre und soziale Verhältnisse bedingten Disparitäten auszugleichen (vgl. Prenzel et al. 2007). Dabei werden als eine Facette der Bildungsbenachteiligung auch kreative Potenziale vergeudet. Im Rahmen des Teilprojektes sollen nun zum einen (JeKi-spezifische) Möglichkeiten der Integration von Schülerinnen und Schülern mit unterschiedlichem sozio-ökonomischen, familiärem und kulturellem Hintergrund herausgearbeitet werden. Zum anderen steht die Rekonstruktion von Maßnahmen für eine gelingende Ermöglichung kultureller Teilhabe im Zentrum. Weiterhin interessieren mögliche Aus- und Rückwirkungen auf individuelle Einstellungen der JeKi-Grundschul Kinder zum schulischen Lernen insgesamt sowie auf das Elternhaus. Dabei sollen regionale und konzeptionelle Unterschiede berücksichtigt werden.

Das Untersuchungsdesign sieht qualitative Fallstudien an fünf ausgewählten Grundschulen (zwei in Hamburg, drei im Ruhrgebiet) zu zwei Erhebungszeitpunkten vor. Dabei wird sich die kriterienbezogene Auswahl der gegensätzlichen Schulen auf Aspekte wie die Lage der Schule in ihrem Umfeld, den prozentualen Anteil von Schülerinnen und Schülern mit Migrationshintergrund, den sozioökonomischen Hintergrund der Eltern oder auch konzeptionelle Schwerpunkte stützen. Die Datenerhebung mittels Gruppeninterviews mit Schülerinnen und Schülern, Eltern, dem pädagogischen Personal sowie der Schulleitung bilden die empirische Grundlage. Zusätzlich dazu soll die Entwicklung der musikalischen Praxis von Kindern und Eltern beobachtet werden: zum einen werden die Kinder regelmäßig Bilder von sich und ihren musikalischen Erlebnissen malen. Zum anderen können die Eltern parallel dazu wahlweise ein vorbereitetes Tagebuch führen oder per Foto (herkömmlich, digital oder als MMS) die musikalischen Aktionen ihrer Kinder dokumentieren. Auf diese Weise können Schwerpunkte des Teilhabekonstrukts (wie Interaktion, zeitlicher Verlauf, biografische Muster) abgebildet und rekonstruiert werden. Schließlich sollen ausgewählte Dokumente wie bspw. schulspezifische

JeKi-Konzeptionen analysiert werden. Mit Verfahren der Triangulation werden sich die verschiedenen Daten multiperspektivisch ergänzen. Zusammenfassend kann formuliert werden, dass JeKi als musikalisches Grundschulkonzept zunächst *Zugangschancen* zu kultureller Teilhabe im schulischen Kontext birgt. Allerdings geht es um mehr, was im Rahmen des Teilprojektes untersucht werden soll:

„Kulturelle Teilhabe muss nämlich nicht nur überprüfen, ob bei einem bestimmten Angebot die [...] Bedingungen (Erreichbarkeit, Finanzierbarkeit, bildungsmäßige Voraussetzungen) erfüllt sind. Man muss für eine ständige, kontinuierliche Angebots- und Beteiligungsstruktur sorgen, damit man in (kultureller) Übung bleibt. [...] Man braucht nämlich Infrastrukturen, die auf Dauer versuchen, mit entsprechenden Angeboten auch in schlecht zugänglichen Gegenden zurückhaltende Zielgruppen zu erreichen.“ (Max Fuchs 2008, S. 6)

E Teilprojekt: Musikalische Präferenzentwicklung bei Grundschulkindern

Veronika Busch & Julika Wagschal (Universität Bremen)

Das SIGrun-Teilprojekt ‚Präferenz‘ untersucht die Auswirkungen vertieften Musikunterrichts im Rahmen von JeKi auf die musikalische Präferenzentwicklung von Grundschulkindern. Den Forschungshintergrund bildet die aktuelle musikpsychologische und musikpädagogische Diskussion zum Konzept der Offenohrigkeit („open-earedness“; Hargreaves 1982a, 1982b), mit dem die positive Beurteilung vielfältiger Musikstile von Kindern zu Beginn der Grundschulzeit beschrieben wird. Die bisherige Forschungslage zur Offenohrigkeit zeigt, dass es ab einem Alter von etwa neun Jahren zu einer Einengung der Musikpräferenz auf Pop- und Rockmusik kommt. Hierbei äußern Jungen früher als Mädchen deutliche Musikpräferenzen und -abneigungen (Gembris & Schellberg 2003; Kopiez & Lehmann 2008).

Mittels quantitativer und qualitativer Methoden wird im Rahmen einer quasiexperimentellen kontrollierten Längsschnittstudie der Frage nachgegangen, wie vertiefter Musikunterricht die musikalische Präferenzentwicklung beeinflusst und was die Gründe für die beobachtete Entwicklung sind. Hierbei wird die Musikpräferenz von Grundschulkindern mit und Grundschulkindern ohne vertieften Musikunterricht im Verlauf der ersten vier Schuljahre verglichen und auf folgende potentielle Einflussfaktoren fokussiert: u. a. Persönlich-

keitsmerkmale, Hörgewohnheiten im Alltag, musikalisches Erleben und Verhalten sowie musikalische Toleranz.

In Kooperation mit dem SIGrun-Teilprojekt ‚Kulturelle Teilhabe‘ wird zudem untersucht, inwieweit eine – möglicherweise durch JeKi gestärkte – musikalische Erfahrungsoffenheit die Entwicklung einer generellen kulturellen Offenheit begünstigt.

Die Gesamtstichprobe setzt sich aus zwei Teilstichproben zusammen: 1. Grundschulkinder von Schulen mit einem Förderschwerpunkt im Bereich Musik (11 JeKi-Schulen; N= 397) und 2. Grundschulkinder von Schulen mit einem Förderschwerpunkt im Bereich Sport (6 Sportschulen; N=243). Die Kinder hören zwölf Musikbeispiele, die sich hinsichtlich Instrumentation, Genre und kulturellem Kontext unterscheiden und bewerten diese auf einer fünfstufigen ikonographisch verankerten Skala (Smileys von 1 „will ich unbedingt hören“ bis 5 „will ich absolut nicht hören“). Zudem beantworten die Grundschulkinder einen Fragebogen zur Toleranz im Umgang mit Musik. Des Weiteren wird ein Persönlichkeitsfragebogen (Fünf-Faktoren-Fragebogen für Kinder; Asendorpf 2003) eingesetzt, in dem die Eltern Fragen zur Persönlichkeit ihres Kindes beantworten. Für die qualitativen Erhebungen wird aus der Gesamtstichprobe eine Gruppe von 30 Kindern aus fünf Schulen ausgewählt, mit denen Gruppeninterviews zu den Gründen der musikalischen Präferenz durchgeführt werden. Zudem werden Eltern und Lehrkräfte hinsichtlich der Gründe für die Entwicklung der kindlichen Musikpräferenz befragt. Die quantitativen Erhebungen finden jeweils am Ende des ersten, zweiten, dritten und vierten Grundschuljahres statt. Die erste Datenerhebung wurde bereits am Ende des Schuljahres 2008/09 durchgeführt. Die qualitativen Erhebungen finden jeweils zur Mitte der zweiten und vierten Grundschulklasse statt.

5 Forschungsprojekt „Co-Teaching im JeKi-Unterricht“

Erich Beckers (Technische Universität Braunschweig)

Das gemeinsame Unterrichten von Grundschullehrern und Musikschullehrern⁴ mit entsprechend unterschiedlicher pädagogischer und musikalischer Ausbil-

4 Die maskuline Schreibweise wird aus Gründen der besseren Lesbarkeit des Textes durchgängig verwendet und inkludiert immer Lehrerinnen und Lehrer bzw. Schülerinnen und Schüler.

dung ist ein besonderes Merkmal des Programms „Jedem Kind ein Instrument“. In NRW soll diese Form des Co-Teaching⁵ in allen Gruppen des 1. Unterrichtsjahres praktiziert und damit flächendeckend eingeführt werden. Über die genauen Bedingungen des Co-Teaching, dessen Möglichkeiten und Grenzen, Methoden und Wirkungen im JeKi-Unterricht ist jedoch noch wenig bekannt. Co-Teaching wurde bislang insgesamt wenig beforscht. Zu nennen wären exemplarisch Frauke Janz (2006), die in ihrer Dissertation die „Interprofessionelle Kooperation in Klassenteams von Schülerinnen und Schülern mit schweren und mehrfachen Behinderungen“ untersucht oder auch Brigitte Frommherz und Therese Halfhide (2003) mit ihren Untersuchungen zum Teamteaching an Unterstufen-Schulklassen der Stadt Zürich. Der Ertrag dieser Studien für das geplante Forschungsprojekt ist jedoch eher begrenzt, zumal diese Arbeiten keinen musikpädagogischen Bezug aufweisen und zudem methodologisch eher an der Oberfläche bleiben.

Verwertbare Anhaltspunkte zum Problem des Co-Teaching im Musikunterricht der Grundschule liefert hingegen die Evaluationsstudie zum Bochumer JeKi-Projekt (Beckers & Beckers 2008). Unter Berücksichtigung der dort gefundenen Evaluationsergebnisse lässt sich die Fragestellung des Forschungsvorhabens wie folgt entfalten: Unter welchen institutionellen und individuellen Bedingungen und Voraussetzungen findet Co-Teaching im JeKi-Unterricht statt und wie wirken sich die unterschiedlichen Bedingungen und Voraussetzungen und die damit einhergehenden heterogenen Umsetzungen des Co-Teaching auf den Unterricht und auf das Lernen der Kinder aus?

Das Forschungsprojekt folgt einer im Grundsatz eher qualitativen Forschungslogik i. S. einer qualitativen Systemanalyse und beinhaltet Lehrerbefragungen, Unterrichtsbeobachtungen und Schülerbefragungen. Die jeweiligen Bereiche werden dabei nicht isoliert betrachtet, sondern in ihrer Wechselwirkung zueinander. Daher sind die Untersuchungen so angelegt, dass die Ergebnisse direkt aufeinander bezogen werden können und im Sinne einer Triangulation der gegenseitigen Ergänzung und Absicherung dienen.

Die Co-Teaching-Bedingungen werden im Rahmen einer zweiphasigen Lehrerbefragung erfasst, in der ersten quantitativ, darauf aufbauend – in der zweiten – qualitativ. In der ersten Phase sollen mittels Fragebögen die verschiedenen Tandemkonstellationen ermittelt werden. Berücksichtigt werden

5 Gebräuchlich sind weitere Begriffe wie Team- oder Tandem-Teaching etc. Eine differenzierte Begriffsdiskussion erfolgt im Rahmen des Forschungsprojekts.

hier u. a. die musikalischen und pädagogischen Voraussetzungen, die die Co-Lehrer mitbringen (z. B. Grundschullehrer mit/ohne musikalische/r Ausbildung/Instrumentalausbildung und Musikpädagogen mit/ohne EMP-Erfahrung⁶) und die Bedingungen, unter denen in den Grundschulen unterrichtet wird (z. B. Lage u. Ausstattung der Schulen, soziale Lage der Schüler etc.). Anhand der gefundenen Daten werden Cluster gebildet, auf deren Grundlage eine Klassifizierung der unterschiedlichen Tandemtypen stattfindet.

Die 2. Phase der Lehrerbefragung erfolgt als qualitative Strukturanalyse, indem von denjenigen Lehrern, die als typische Repräsentanten der jeweiligen Cluster ermittelt wurden, Subjektive Theorien über Co-Teaching im JeKi-Unterricht rekonstruiert werden.⁷ Diese Subjektiven Theoriestrukturen sollen einen vertiefenden Einblick gewähren in die handlungsleitenden Erfahrungen, Einstellungen und Erwartungen der Lehrer bezüglich des Co-Teaching. Zu diesem Zweck werden die betreffenden Grundschul- und Musikschullehrer zunächst einem qualitativen Einzelinterview unterzogen. Ausgehend von den Interviewergebnissen werden dann gemeinsam mit den Lehrern ihre subjektiven Theoriestrukturen über Co-Teaching rekonstruiert und anschließend kommunikativ validiert.

Es folgen systematische Unterrichtsbeobachtungen zur Frage nach möglichen Wirkungen des Co-Teaching auf den Unterricht und auf das Schülerverhalten. Dokumentiert werden sollen u. a. Merkmale der Zusammenarbeit der beiden Co-Lehrer, Aspekte des Unterrichts, das Lernverhalten der Schüler und die Sozialbeziehungen unter Berücksichtigung der Tandemkonstellationen und möglicher Gender-Effekte. Es sollen genau diejenigen Unterrichtsgruppen beobachtet werden, von deren Lehrern subjektive Theorien über Co-Teaching vorliegen. Die Ergebnisse der Unterrichtsbeobachtungen werden dann zu den Fragebogenergebnissen und den Subjektiven Theoriestrukturen in Beziehung gesetzt und daraufhin befragt, wie sich die verschiedenen prototypischen Co-Teaching-Bedingungen und -konstellationen auf den JeKi-Unterricht und auf die daran beteiligten Personen auswirken. Vergleichend sollen zudem Unterrichtsbeobachtungen in Gruppen durchgeführt werden, die nicht im Co-Teaching unterrichtet werden mit dem Ziel, etwas darüber zu erfahren, ob

6 EMP = Elementare Musikpädagogik

7 Zum Konstrukt der Subjektiven Theorien vgl. u. a.: Groeben et al. 1988; Beckers 2004; Niessen 2006.

bzw. worin sich der JeKi-Unterricht sowie das Lernverhalten der Kinder in Gruppen mit und ohne Co-Teaching unterscheiden.

Die Erfahrungen der Kinder mit dem Co-Teaching sowie ihre Sichtweisen und Einstellungen werden im Rahmen von Schülerbefragungen ermittelt. Für diesen Erhebungsschritt wird ein altersangemessenes Fragespiel entwickelt, das Erkenntnisse zu der Frage liefern soll, wie sich Co-Teaching aus Sicht der Schüler auf den Unterricht auswirkt. Da auch hier wieder diejenigen Schüler befragt werden, deren Unterricht beobachtet wurde bzw. von deren Lehrern subjektive Theoriestrukturen zum Co-Teaching vorliegen, sollen diese Ergebnisse nicht nur zu einer multiperspektivischen Problemsicht, sondern auch zur wechselseitigen Validierung der einzelnen Erhebungsschritte beitragen.

Neben dem erhofften Erkenntnisgewinn im untersuchten Problemfeld soll das Forschungsprojekt einen Beitrag zur Qualitätsentwicklung und -sicherung des JeKi-Unterrichts leisten und darüber hinaus die empirische Basis für die Entwicklung eines didaktischen Konzepts zum Co-Teaching im Bereich der Musikvermittlung liefern.

6 Forschungsprojekt BEGIn: Bielefelder Evaluationsstudie zum Gruppen-Instrumentalunterricht

Ulrike Kranefeld, Julia Rheingans, Manfred Nusseck, Timo Fischinger (Universität Bielefeld)

Die Bielefelder Evaluationsstudie BEGIn richtet sich auf zwei zentrale Aspekte der JeKi-Programme: (1) auf die Teilnahme und den Verbleib im Programm und (2) auf die Voraussetzungen und Bedingungen des Gruppeninstrumentalunterrichts in der Grundschule.

Ein wesentliches Ziel der Untersuchung ist die Bestimmung von Prädiktoren für „Drop-Out“, d. h. den freiwilligen Abbruch des Programms durch die Teilnehmer.⁸ In der Literatur gibt es Hinweise darauf, dass Faktoren wie das Übe-Verhalten, das Unterstützungsverhalten der Eltern und das Selbstkonzept Einflüsse auf das Fortsetzen des Instrumentalspiels haben (Hallam 1998). Schon Frakes (1984) zeigte, dass Abbrecher sich selbst als weniger musika-

8 Die Drop-Out-Studie wird nur in NRW durchgeführt, da das JeKi-Programm in den beteiligten Grundschulen in Hamburg für alle Schüler dieser Schulen verbindlich und kein freiwilliges Angebot ist.

lich bezeichneten und weniger familiäre Unterstützung erhielten, wohingegen eine positive Selbstwahrnehmung der musikalischen Fähigkeiten mit einem freiwilligen Weitermachen assoziiert war. Gerade auch die Rolle der Eltern erscheint für die musikalische Entwicklung des Kindes von großer Bedeutung: Faktoren wie die eigene musikalische Erfahrung, persönliches Interesse und das Unterstützungsverhalten spielen dabei eine Rolle (McPherson 2008, Chretech & Hallam 2009).

Neben der Frage nach den Gründen für Drop-Out widmet sich die Studie außerdem der Nachhaltigkeit der JeKi-Programme in Bezug auf die Fortführung des Instrumentalspiels nach Beendigung des 4. Schuljahres. Im Rahmen einer geschichteten Panelstichprobe nehmen 1200 Schülerinnen und Schüler in NRW und ca. 200 in Hamburg sowie deren Eltern an jährlichen Fragebogenerhebungen von der 2. bis 5. Klasse teil.

Ebenso wird die Perspektive der Instrumentallehrenden erfasst, indem sie zu ihren Einstellungen und Erfahrungen mit dem Gruppeninstrumentalunterricht in der Grundschule befragt werden. Eine ergänzende Befragung der Klassenlehrerinnen- und -lehrer und der Schulleitungen zielt darauf, den Stellenwert zu bestimmen, der dem JeKi-Programm innerhalb der Einzelschule zugeschrieben wird.

Um nicht allein die Voraussetzungen für und Einstellungen zum Instrumentalunterricht zu erfassen, wird in einem weiteren Schritt der Instrumentalunterricht selbst in den Blick genommen: Einzelne Unterrichtsstunden in 10-15 Instrumentalgruppen in NRW und in 3-5 Gruppen in Hamburg werden einmal im Jahr per Videokamera aufgezeichnet und mit Hilfe quantifizierender Methoden der Videoanalyse ausgewertet. Wie etwa die Videostudie innerhalb der TIMSS-Studie zum Mathematikunterricht zeigt, ermöglicht die videobasierte Unterrichtsforschung eine Fokussierung auf bestimmte Untersuchungsbereiche, wie z. B. die Identifikation typischer „Inszenierungsmuster“ (Hugener, Pauli & Reusser 2007), in unserem Fall die des Gruppeninstrumentalunterrichts unter den spezifischen Bedingungen der Grundschule. Hierzu werden in einer explorativen Phase mit Hilfe qualitativer Datenanalyse Kategorien für eine quantifizierende Videoanalyse entwickelt, etwa für den Bereich des Umgangs mit Heterogenität in der Lerngruppe.

Die mehrperspektivische Anlage der Untersuchung orientiert sich an der Idee eines Angebot-Nutzen-Modells zur Unterrichtsqualität, wie es Reusser & Pauli in Anlehnung an Modelle von Fend, Helmke und Weinert entwickeln (Reusser & Pauli 2003; Fend 1981; Helmke & Weinert 1997; Helmke 2006).

Im Zentrum des Modells stehen der Unterricht und die Vorstellung, dass in ihm ein Lehrangebot (mit einer spezifischen Quantität und Qualität) auf eine entsprechende Nutzung trifft. Während in der Fragebogenerhebung sowohl die sogenannten angebotsbezogenen wie die nutzungsbezogenen Stützsysteime im Mittelpunkt stehen, richtet sich die Videographie auf den Unterrichtsprozess selbst.

Zur Angebotsseite gehören z. B. Konzepte zum JeKi-Unterricht, pädagogische Traditionen des Gruppeninstrumentalunterrichts, der Stellenwert von JeKi an der Einzelschule oder Lehrer-Merkmale wie ihre musikalische Ausbildung, Einstellungen und subjektive Theorien zu JeKi. Auf der Seite der Nutzung stehen z. B. Variablen wie der soziokulturelle Kontext der Schüler, das Unterstützungsverhalten der Eltern und weitere Schüler-Merkmale wie das musikalische Selbstkonzept oder die Übe-Motivation.

Im Forschungsprojekt BEGIn wird der JeKi-Unterricht also in einem Bedingungsgefüge kontextualisiert, wobei Aspekte des Jeki-Angebots und seiner Nutzung aufeinander bezogen werden. Durch videobasierte Unterrichtsfor-schung wird zudem der Gruppeninstrumentalunterricht im Hinblick auf die besonderen Bedingungen in der Lernumgebung Grundschule erschlossen.

Die Erkenntnisse, die in dieser Evaluationsstudie gewonnen werden, können als empirische Basis zur Weiterentwicklung der JeKi-Programme in NRW und Hamburg genutzt werden.

7 Datenpool des JeKi-Forschungsschwerpunkts

Andreas Lehmann-Wermser & Klaudia Schulte (Universität Bremen)

Ein Teil des Gesamtkonzeptes der JeKi-Begleitforschung ist die Stelle des so genannten ‚zentralen Datenpools‘. Ihre Existenz verdankt sie einer strategischen Überlegung des BMBF, die für die Musikpädagogik eine neue Qualität des Forschungsprofils belegt. Ursprünglich war diese Stelle nur für das Projekt SIGrun geplant (s. die Ausführungen weiter oben). Die Ausweitung als zentrale Einheit für alle Projekte ist verbunden mit zwei Zielen.

Zum einen wurde in der Planungsphase der Erhebungen koordiniert, dass gleiche Kernfragen in den Projekten verwendet werden. Welche dies verbindlich sind und in welcher Form sie erhoben werden, wurde durch Absprachen unter den Wissenschaftlern geklärt. Der Vorteil dieser Kernfragen, beispielsweise gleicher Fragen zum Sozialstatus der Familien von Schülerinnen und

Schülern, liegt darin, dass ausgewählte Variablen über die einzelnen Stichproben der Forschungsverbünde und -projekte verglichen werden können. Die Frage der Repräsentativität etwa lässt sich damit leichter beantworten. Weitergehende statistische Verfahren (etwa des „propensity score matching“) erlauben es theoretisch, auf der ungleich größeren Datenbasis Probleme fehlender Werte oder des längsschnittlichen Designs anzugehen. Zum anderen sollen alle Daten archiviert werden und für spätere Sekundäranalysen zugänglich sein. Diese Idee entspringt der Erfahrung, dass in der Vergangenheit mehrfach durch öffentliche Gelder finanzierte umfangreiche Erhebungen durchgeführt wurden, deren Daten dann von einzelnen Wissenschaftler/-innen unter Verschluss gehalten wurden. Dies mag teilweise verständlich sein, jedoch wird damit zugleich ein potentieller Erkenntnisgewinn verspielt. In der JeKi-Begleitforschung haben sich, um diesem Problem vorzubeugen, alle Projekte verpflichtet, ihre Daten mittel- und langfristig an den Datenpool zu übergeben. Dort werden sie systematisch archiviert, nach Ablauf der Förderungsdauer in eine größere Datenbank eines Forschungsinstituts integriert und damit für Sekundäranalysen bereitgestellt. Formale Regelungen zwischen den Instituten garantieren, dass Qualifikationsarbeiten geschützt bleiben und keine Überschneidungen zwischen Fragestellungen auftreten.

Literatur

- Altenmüller, E. (2006): Neuronale Auswirkungen musikalischen Lernens im Kindes- und Jugendalter und Transfereffekte auf Intelligenzleistungen. In: Schumacher, R. (Hrsg.): *Macht Mozart schlau*, Bonn: BMBF.
- Arbeitsgruppe Deutsche Child Behavior Checklist (1998): Elternfragebogen über das Verhalten von Kindern und Jugendlichen; deutsche Bearbeitung der Child Behavior Checklist (CBCL/4-18). Einführung und Anleitung zur Handauswertung mit deutschen Normen, bearbeitet von M. Döpfner, J. Plück, S. Bölte, K. Lenz, P. Melchers & K. Heim. (2. Aufl.). Köln: Arbeitsgruppe Kinder-, Jugend- und Familiendiagnostik (KJFD).
- Arnoldt, B. (2007): Öffnung von Ganztagsschule. In: Holtappels, H. G.; Klie-me, E.; Rauschenbach, T.; Stecher, L. (Hrsg.): *Ganztagsschule in Deutschland. Ergebnisse der Ausgangserhebung der "Studie zur Entwicklung von Ganztagsschulen" (StEG)*. Weinheim und München: Juventa (Studien zur ganztägigen Bildung), S. 86–105.

- Asendorpf, J. B. & van Aken, M. A. G. (2003). Validity of Big Five Personality Judgments in Childhood: A 9 Year Longitudinal Study. *European Journal of Personality*, 17, 1–17.
- Bartelheimer, P. (2008): Was bedeutet Teilhabe. Risiken und Nebenwirkungen. In: *Kulturelle Bildung*, Jg. 2, H. 2, S. 21-23.
- Bastian, H. G. (2000). *Musik(erziehung) und ihre Wirkung*. Mainz : Schott.
- Bastian, H. G. (2001). *Kinder optimal fördern - mit Musik. Intelligenz, Sozialverhalten und gute Schulleistungen durch Musikerziehung*. Mainz: Schott.
- Beckers, E. & Beckers, R. (2008): *Faszination Musikinstrument – Musikmachen motiviert. Bericht über die 2-jährige Evaluationsforschung zum Bochumer Projekt „Jedem Kind ein Instrument“*. Zukunftsstiftung Bildung in der GLS Treuhand e. V. Bochum (Hg.). LIT.
- Beckers, E. (2004): *Erwachsene lernen Musik. Empirische Studien zu subjektiven Theorien des Musikkernens Erwachsener aus Sicht der Lernenden*. LIT.
- Behr-Heintze, A. & Lipski, J. (2005): *Schulkooperationen. Stand und Perspektiven der Zusammenarbeit zwischen Schulen und ihren Partnern*. Schwalbach/Ts.: Wochenschau Verlag.
- Bongard, S., Kelava, A., Sabic, M., Aazami-Gilan, D. & Kim, Y.B. (2007): *Akkulturation und gesundheitliche Beschwerden bei drei Migrantenstichproben in Deutschland*. In: H. Eschenbeck, U. Heim-Dreger, & C.W. Kohlmann, (Hrsg.): *Beiträge zur Gesundheitspsychologie* (S. 53). Schwäbisch Gmünd: Gmünder Hochschulreihe Band 29.
- Bonsen, M. et al. (2008): *Bildungsrelevante Ressourcen im Elternhaus: Indikatoren der sozialen Komposition der Schülerschaften an Dortmunder Schulen*. In: Stadt Dortmund. *Der Überbürgermeister* (Hg.): *Erster kommunaler Bildungsbericht für die Schulstadt Dortmund. Schulentwicklung Dortmund*. Münster: Waxmann, S. 125-149.
- Brickenkamp, R. (2002): *Aufmerksamkeits-Belastungs-Test*. 9. überarbeitete und neu normierte Auflage. Göttingen: Hogrefe.
- Byo, J. (1991): *An assessment of musical instrument preferences of third-grade children*. In: *Bulletin of the Council for Research in Music Education*. H. 110, S. 21-32.
- Chreech, A. & Hallam S. (2009): *Interaction in instrumental learning: the influence of interpersonal dynamics on parents*. *International Journal of Music Education*, 27, 94-106.

- Costa-Giomi, E. (2004): Effects of three years of piano instruction on children's academic achievement, school performance and self-esteem. *Psychology of Music*, 32 (2), 139-152.
- Czerwanski, A. (2003): Netzwerke als Praxisgemeinschaften. In: Czerwanski, A. (Hrsg.): *Schulentwicklung durch Netzwerkarbeit. Erfahrungen aus den Lernnetzwerken im "Netzwerk innovativer Schulen in Deutschland"*. (S. 9–15). Gütersloh: Bertelsmann-Stiftung.
- Ditton, H. (2009): Evaluation und Qualitätssicherung. In: Tippelt, R.; Schmidt, B. (Hg): *Handbuch Bildungsforschung*. 2. überarbeitete und erweiterte Auflage. Wiesbaden: Verlag für Sozialwissenschaften, S. 607-623.
- Esslinger, I. (2002): *Berufsverständnis und Schulentwicklung: ein Passungsverhältnis? Eine empirische Untersuchung zu schulentwicklungsrelevanten Berufsauffassungen von Lehrerinnen und Lehrern*. Bad Heilbrunn: Julius Klinkhardt.
- Fend, H. (1981): *Theorie der Schule* (2. Aufl.). München: Urban & Schwarzenberg.
- Frakes, L. (1984): Differences in music achievement, academic achievement and attitude among participants, dropouts and non-participants in secondary school music. Ph.D. Thesis, University of Iowa, Dissertation Abstracts International, 46(02), 370A.
- Frommherz, B. & Halfide, Th. (2003): *Teamteaching an Unterstufenklassen der Stadt Zürich*. Pädagogisches Institut der Universität Zürich. http://www.ife.uzh.ch/pp1/stoeckli/Projektarchiv_files/TeamteachingZH.pdf.
- Fuchs, M. (2008): Kulturelle Bildung und Teilhabe. In: *Kulturelle Bildung*, Jg. 2, H. 2, S. 5-6.
- Fuchs, M. & Bockhorst, H. (2008): *Kultur - Teilhabe - Bildung. Reflexionen und Impulse aus 20 Jahren*. München: kopaed.
- Gembris, H. & Schellberg, G. (2003): Musical preferences of elementary school children. Paper presented at the 5th Triennial Conference of the European Society for the Cognitive Sciences of Music (ESCOM, 8.-13.9.2003). Hannover.
- Glaser, B. G.; Strauss, A. L. (1998): *Grounded theory. Strategien qualitativer Forschung*. Bern: Huber.
- Gordon, E. E. (1986): *Intermediate measures of music audiation*. Chicago: GIA.

- Groeben, N., Wahl, D., Schlee, J. & Scheele, B. (1988): Forschungsprogramm Subjektive Theorien. Eine Einführung in die Psychologie des reflexiven Subjekts. Francke.
- Hämäläinen, J. A., Salminen, H. K. & Leppänen, P. H. T. (in press): Basic auditory processing deficits in dyslexia – Review of the behavioral and event-related potential/field evidence. *Journal of Learning Disabilities*.
- Hallam, S. (1998): The Predictors of Achievement and Dropout in Instrumental Tuition. *Psychology of Music*, 26, 116-132.
- Hargreaves, D. J. (1982): Preference and prejudice in music: A psychological approach. *Popular Music and Society*, 8, 13-18.
- Hargreaves, D. J. (1982): The development of aesthetic reactions to music. *Psychology of Music (Special Issue)*, 51-54.
- Heller, K. & Geisler, H. J. (1983): Kognitiver Fähigkeitstest für 1. bis 3. Klassen. Göttingen: Hogrefe.
- Helmholtz, H. L. F. v. (1885). On the sensations of tone. London: Longmans.
- Helmke, A. (2006): Unterrichtsqualität: Erfassen, Bewerten, Verbessern (4. Aufl.). Seelze: Kallmeyersche Verlagsbuchhandlung.
- Helmke, A. & Weinert, F. E. (1997): Bedingungsfaktoren schulischer Leistungen. In F. E. Weinert (Hrsg.), *Psychologie des Unterrichts und der Schule (Enzyklopädie der Psychologie, Pädagogische Psychologie, Vol. 3, S. 71-176)*. Göttingen: Hogrefe.
- Helms, S. (2002): Allgemein bildende Schule und Musikschule. In: Helms, S. (Hrsg.): *Allgemein bildende Schule und Musikschule in europäischen Ländern*. Kassel: Gustav Bosse (Musik im Diskurs, 17), S. 7–33.
- Helmstaedter, C., Lendt, M. & Lux, S. (2001): Verbaler Lern- und Merkfähigkeitstest. Göttingen: Hogrefe.
- Hugener, I., Pauli, C. & Reusser, K. (2007): Inszenierungsmuster, kognitive Aktivierung und Leistung im Mathematikunterricht. Analysen aus der schweizerisch-deutschen Videostudie. In D. Lemmermöhle, M. Rothgangel, S. Bögeholz, M. Hasselhorn & R. Watermann (Eds.), *Professionell Lehren - Erfolgreich Lernen*. Münster: Waxmann, 109-121.
- Irion, Th. (2008): Hypermedia-Recherche im Grundschulalter. Eine qualitative Videostudie zu Vorerfahrungen und Recherchekompetenzen. Baltmannsweiler: Schneider Hohengehren.

- Jansen, D. (2006): Einführung in die Netzwerkanalyse: Grundlagen, Methoden, Forschungsbeispiele (3., überarb. Aufl.). Wiesbaden: VS Verl. für Sozialwissenschaften.
- Jäncke, L. (2008). Macht Musik schlau? Basel: Huber.
- Jäncke, L., Shah, N., Posse, S., Grosse-Ryken, M., & Müller-Gärtner, H. (1998): Intensity coding of auditory stimuli: an fMRI study. *Neuropsychologia*, 36(9), 875-883.
- Janz, F. (2006): Interprofessionelle Kooperation in Klassenteams von Schülerinnen und Schülern mit schweren und mehrfachen Behinderungen. Eine empirische Untersuchung in Baden-Württemberg.
- Jerusalem, M. & Schwarzer, R. (1991). Entwicklung des Selbstkonzepts in verschiedenen Lernumwelten. In: Pekrun, R. & Fend, H. (Hrsg.): *Der Mensch als soziales und personales Wesen: Bd. 11. Schule und Persönlichkeitsentwicklung. Ein Resümee der Längsschnittforschung.* (S. 115–128). Stuttgart: Ferdinand Enke.
- Kopiez, R. & Lehmann, M. (2008): The ‘open-earedness’ hypothesis and the development of age-related aesthetic reactions to music in elementary school children. *British Journal of Music Education*, 25(2), 121–138.
- Kranefeld, U. (2008): Zur Standortbestimmung einer Interpretativen Unterrichtsforschung in der Musikpädagogik im Spannungsfeld von erziehungswissenschaftlicher und fachdidaktischer Perspektive. In: Jürgen Vogt et al. (Hrsg.): *Musikpädagogik auf dem Weg zur Vermittlungswissenschaft? Sitzungsbericht 2007 der Wissenschaftlichen Sozietät Musikpädagogik (= Wissenschaftliche Musikpädagogik; 2).* Münster: Lit, S. 99-111.
- Kunze, I. (2008): Begründungen und Problembereiche individueller Förderung in der Schule. Vorüberlegungen zu einer empirischen Untersuchung. In: Kunze, I. & Solzbacher, C. (Hg.): *Individuelle Förderung in der Sekundarstufe I und II.* Baltmannsweiler: Schneider Hohengehren, S. 13–25.
- Lohaus, A., Eschenbeck, H., Kohlmann, C. W. & Klein-Heßling, J. (2006): Fragebogen zur Erhebung von Stress und Stressbewältigung im Kindes- und Jugendalter. Göttingen: Hogrefe.
- McPherson, G. (2008): The role of parents in children’s musical development. *Psychology of Music*, 37, 91-110.

- Meyer-Clemens, A. (2007): Kooperation zwischen allgemein bildender Schule und Musikschule: Theorie und Praxis – Bedingungen – Evaluation. Marburg: Tectum.
- Nantais, K. M. & Schellenberg, E. G. (1999): The Mozart effect: An artifact of preference? *Psychological Science*, 10, 370-373.
- Niessen, A. (2006): Individualkonzepte von Musiklehrern. (= Theorie und Praxis der Musikvermittlung, 6). Münster: Lit.
- Niessen, A. & Lehmann-Wermser, A. (2006): Musikunterricht im Spiegel mehrperspektivischer Sinnzuschreibungen. In: Knolle, N. (Hg.): Lehr- und Lernforschung in der Musikpädagogik. (= Musikpädagogische Forschung, 27). Essen: Die Blaue Eule. S. 239-252.
- Reuter, C. (1995): Der Einschwingvorgang nichtperkussiver Musikinstrumente. Frankfurt: Lang.
- Oswald, W. D. & Roth, E. (1987): Der Zahlen-Verbindungs-Test. 2., überarbeitete und erweiterte Auflage. Göttingen: Hogrefe.
- Overy, K., Norton A.C., Cronin, K.T., Gaab, N., Alsop, D.C., Winner, E. & Schlaug, G. (2004): Imaging melody and rhythm processing in young children. *Neuroreport*, 15, 1723-1726.
- Prenzel, M.; Artelt, C.; Baumert, J.; Blum, W.; Hammann, M.; Klieme, E. & Pekrun, R. (Hrsg.) (2007): PISA 2006. Die Ergebnisse der dritten internationalen Vergleichsstudie. Münster: Waxmann.
- Rauer, W. & Schuck, K.D. (2003): Fragebogen zur Erfassung emotionaler und sozialer Schulerfahrungen von Grundschulkindern dritter und vierter Klassen. Göttingen: Hogrefe.
- Rauscher, F. H., Shaw, G. L. & Ky, K. N. (1993): Music and spatial task performance. *Nature*, 365, 611.
- Reusser, K. & Pauli, C. (2003): Mathematikunterricht in der Schweiz und in weiteren sechs Ländern. Bericht über die Ergebnisse einer internationalen und schweizerischen Video-Unterrichtsstudie, CD-ROM des Pädagogischen Instituts Zürich.
- Schellenberg, E. G. (2003): Music Lessons Enhance IQ. *Psychological Science*, 15 (8), 511-514.
- Schneider, P., Sluming, V., Roberts, N., Scherg, M., Goebel, R., Specht, H. J., et al. (2005): Structural and functional asymmetry of lateral Heschl's gyrus reflects pitch perception preference. *Nat. Neurosci.*, 8(9), 1241-1247

- Schneider, P & Wengenroth, M. (2009): The Neural basis of individual holistic and spectral sound perception. *Contemporary Music Review*, 28, Nr. 3, S. 315-328.
- Schulten, M. L. & Lothwesen, K. S. (2009): MoMo verbindet! Musik erleben und lernen in der "Musikschule für alle": Abschlussbericht der wissenschaftlichen Evaluation zum Programm "Monheimer Modell - Musikschule für alle". Frankfurt am Main: Linden.
- Schwippert, K., Bos, W. & Lankes, E.-M. (2003): Heterogenität und Chancengleichheit am Ende der vierten Jahrgangsstufe im internationalen Vergleich. In: W. Bos et al. (Hrsg.), *Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich* (S. 265–289). Münster: Waxmann.
- Seither-Preisler, A., Johnson, L., Krumbholz, K., Nobbe, A., Patterson, R., Seither, S., et al. (2007): Tone sequences with conflicting fundamental pitch and timbre changes are heard differently by musicians and nonmusicians. *Journal of Experimental Psychology: Human Perception and Performance*, 33(3), 743-751.
- Seither-Preisler, A., Johnson, L., Seither, S., & Lütkenhöner, B. (2008): The perception of dual aspect tone sequences changes with stimulus exposure. *Brain Research Journal*, 2(3).
- Spychiger, M. (2006): Ansätze zur Erklärung der kognitiven Effekte musikalischer Betätigung. In: R. Schuhmacher (Hrsg.): *Macht Mozart schlau?* Berlin: BMBF.
- Stern, E. (2006): Intelligenz und ihre Entwicklung. In: R. Schuhmacher (Hrsg.): *Macht Mozart schlau?* Berlin: BMBF. S. 131-138.
- Weber, E. W., Spychiger, M. & Patry, J.L. (1993): *Musik macht Schule. Biografie und Ergebnisse eines Schulversuchs mit erweitertem Musikunterricht*. Essen: Die Blaue Eule.
- Weber, E. W., Spychiger, M. & Patry, J.-L. (1993): *Musik in der Schule: Ein ehrgeiziges Projekt*.
Verfügbar unter: http://www.ewaweber.ch/musik_macht_schule.htm [26.5.2009].
- Zatorre, R. J. (2005): Music, the food of Neuroscience. *Nature* 434, 312-315.

Handlungskompetenzen und Studiumsmotivation von Musikstudierenden

1 Einleitung

Die bisherigen Studien im Rahmen der erziehungswissenschaftlichen Lehrerbildungsforschung fokussieren vor allem die Bereiche des Selbstkonzepts und des Kompetenzerwerbs von bereits in der Lehre tätigen Personen. Dieses Forschungsprogramm impliziert, dass interindividuelle Unterschiede bezüglich Erfahrung und Entwicklung der Personen erst durch die Ausübung des Lehrberufes entstanden sind. In wie weit die bestehenden Professionalisierungs- und Individualkonzepte auf das vorangegangene Studium und/oder auf bestehende Persönlichkeits- bzw. Selbstkonzeptdeterminanten zurückzuführen sind, ist bisher nicht hinreichend geklärt. Dieses offenbart sich auch in den vorliegenden Studien seitens der Musikpädagogik. Hier stehen entsprechend Aspekte der Selbstprofessionalisierung und des Selbstkonzepts bzw. der Individualkonzepte von Musiklehrpersonen im Mittelpunkt des Interesses (vgl. Niessen, 2006). Nicht-qualitative Untersuchungen zu dem im Studium erfahrenen Kompetenzerwerb sind bislang rar und methodisch ausbaufähig (vgl. Hoffmann 2007, Neuhaus 2008). Dennoch zeigt z.B. die Längsschnittstudie von Rauin & Maier (2007) aus dem Bereich der erziehungswissenschaftlichen Forschung, die Lehramtsstudierende beim Übergang vom Studium in die Berufstätigkeit begleitete, dass ein frühes Burn-out vor allem bei Lehrpersonen zu beobachten ist, die bereits im Studium überfordert und wenig engagiert waren. Der hierbei angewandte typologische Ansatz erlaubt zwar keine differentiellen Aussagen über interindividuelle Unterschiede, aber dennoch betont diese Untersuchung die hohe Bedeutsamkeit des Studiums einerseits und der Persönlichkeit andererseits. Auf dieser Grundlage kann angenommen werden, dass nicht nur die studiumsrelevanten Motivations- und Persönlichkeitsvariablen an sich von Bedeutung sind, sondern dass zudem auch interindividuelle Differenzen dieser Variablen in Anhängigkeit von unterschiedlichen Ausbildungsformen zum Lehrerberuf bestehen dürften.

Ewald Terhart konstatiert in der deutschsprachigen Lehrerforschung einen ab den 1990er Jahren vollzogenen Wandel der Forschungsinteresses bezüglich der Fertigkeiten und Fähigkeiten von Lehrkräfte, die mit einer Abkehr von Schwerpunkten auf „pädagogische[r] Begabung“ und „Praxischock“ und einer Konzentration auf berufsbiographische Aspekte eine kompetenztheoretische Fundierung ermöglicht (Terhart, 2007: 40f). Entsprechend widmen sich jüngere Studien aus dem Bereich der erziehungswissenschaftlichen Lehrerbildungsforschung den Kompetenzmodellen und Risikofaktoren des Lehrerberufs (vgl. Hilligus & Rinkens, 2006; Lüders & Wissinger, 2007). Das von Terhart vorgelegte Modell unterscheidet drei Komponenten, die kognitive, ethisch-moralische und praktische Aspekte beschreiben. Hierbei wird eine kontinuierliche Entwicklung dieser Bereiche angenommen, die in der Formulierung von Kompetenzmodellen zu berücksichtigen sind (Terhart, 2007: 50f). Allerdings ist die Anwendung und Konkretisierung dieser Komponenten für genuin musikpädagogische Lehrkompetenzen zu leisten. Jürgen Vogts Argumentation bezüglich der inhaltlichen Ausdeutung der Begriffe „Professionalisierung“ und „Praxisbezug“ in der Musiklehrerbildung, geht in dieselbe Richtung und unterstreicht ein durch theoretische Fundierung flexibilisiertes Repertoire an Handlungsoptionen für die Unterrichtspraxis (vgl. Vogt 2002). Die Ergebnisse der Langzeitstudie von Udo Rauin und Uwe Maier zeigen, dass Studierende, die bereits Schwierigkeiten im Studium äußerten auch Probleme im Berufsalltag erleben (vgl. Rauin & Maier, 2007). Eine clusteranalytisch begründete Typologie unterscheidet drei Typen anhand der Selbsteinschätzungen des eigenen Studierverhaltens („riskant“, „engagiert“, „pragmatisch“). Da es sich hier jedoch um einen typologischen Ansatz handelt (Aggregation der Daten über Personen), ist eine Berücksichtigung persönlichkeitsrelevanter Merkmale nur bedingt möglich. Zwar werden aufgrund des Forschungsdesigns Veränderungen bzw. Konstanten der Einstellungen und Erwartungen von Lehramtsstudierenden sichtbar und der Übergang in die ersten Berufsjahre anhand dieser Merkmale nachvollziehbar, Motivationen und Motive der Studienwahl und des Studierverhaltens werden aber lediglich deskriptiv dargestellt und nicht zu etwaigen Persönlichkeits- und Selbstkonzeptfaktoren in Beziehung gesetzt.

In der deutschsprachigen musikpädagogischen Forschung wurden vielfach Studien zum Beruf und Berufsbild von Musiklehrerinnen und Musiklehrern an allgemein bildenden Schulen vorgelegt (vgl. u.a. Schulten, 1979; Kraemer, 1991; Niessen, 2006a u. 2006b). Die Perspektive von Musikstudierenden in Lehramtsstudiengängen wurde bisher eher selten thematisiert. Gleiches gilt für den Übergang vom Studium in das Referendariat. Allerdings finden sich be-

reits einige wenige internationale Studien, die jedoch vor allem die berufspraktische Sozialisation (Professionalisierung) der jeweiligen nationalen Ausbildungsbedingungen fokussieren und somit auf die deutsche Situation nur bedingt übertragbar sind (vgl. Bouji, 2004; Roulston, Legett & Womack, 2005). Hans Günther Bastian stellte in einer umfassenden Dokumentation musikpädagogischer Forschung in Deutschland (Bastian & Kraemer, 1992) fest, dass sich lediglich 6% der zwischen 1985-1990 durchgeführten Forschungsprojekte studentischen Belangen widmen und konstatiert in dieser Hinsicht ein „Defizit“, das schon auf die in seiner nachfolgenden Untersuchung (Bastian, 1995:43) aufgestellte Forderung zielt, Ergebnisse derartiger Studien auch für die Ausbildungspraxis nutzbar zu machen. Die bislang vorliegenden Untersuchungen zu Motivationen oder Bedürfnissen von Musikstudierenden weisen in der Regel eine praxisorientierte Perspektive auf, indem das Forschungsinteresse vor allem auf die Veränderungen der Studienbedingungen oder Berufspraxis abzielen. Allerdings liegt der Schwerpunkt der untersuchten Stichproben deutlich auf den Lehramtsstudiengängen und Studierende in künstlerisch orientierten Fächern werden kaum berücksichtigt. Hinzu kommt, dass die bestehenden Studien unterschiedliche Ziele und Perspektiven im Blick haben. So sind z.B. einige direkt bezogen auf hochschulpolitisch begründete (Um)Strukturierungen der musikpädagogischen Ausbildung (vgl. hierzu u.a. die thematisch entsprechend orientierten Bände der Reihe *Forschung in der Musikpädagogik*).

Zusammenfassend lassen sich die in der Literatur publizierten Studien bezüglich folgender Schwerpunkte unterscheiden: Untersuchungen zur Studienwahl (z.B. Darius, 1976; Zimmerschied 1978; Bastian, 1995; Neuhaus, 2008), zur Sozialisation und Studienbedingungen (z.B. Hörmann, 1981; Grimmer, 1986 u. 1987; Stroh, 1996; Siedenburger, 2009), zur Motivation im Studium (z.B. Behne, 1979) und zur Qualifikation durch das Studium (z.B. Hoffmann, 2007). Die musikpädagogischen Studien im deutschsprachigen Raum lassen hierbei folgende Schlüsse zu: (a) Es scheinen studiengangsspezifische Einstellungen und Erwartungen an Studium und Beruf des Musiklehrens an allgemein bildenden Schulen zu existieren, die jedoch bislang nicht auf der Grundlage quantitativ-empirischer Arbeiten der Forschung und Überprüfbarkeit zugänglich gemacht wurden. (b) Vielmehr liegt der Schwerpunkt auf einer praktischen Anwendbarkeit der Ergebnisse in Studienberatung und Lehrerbildung, was zu einer (leider) eher geringen wissenschaftlichen, grundlagenorientierten Diskussion um Theorie und Methoden führt. (c) Ein systematisches Erklärungsmodell bezüglich persönlichkeitsrelevanter Merkmale fehlt. Gabriele

Hoffmann (2007) eröffnet mit der Einbindung des Konzepts der Selbstwirksamkeit einen Weg in diese Richtung. Allerdings werden hier, durch den Ausschluss möglicher berufstypischer Risikofaktoren, noch keine Beziehungen zu möglichen Qualifikationsmerkmalen oder Handlungskompetenzen erfasst.

2 Forschungsfragen

Im Rahmen der vorliegenden Arbeit sollten erstmals Persönlichkeits- und Coping-Variablen in Verbindung mit subjektiven Kompetenz- und Motivationskonstrukten von Musikstudierenden gebracht werden. Dabei standen folgende Forschungsfragen und Hypothesen im Mittelpunkt: (a) Welcher Zusammenhang besteht zwischen Persönlichkeit und der selbstbeschriebenen Kompetenz bzw. Leistungsmotivation? (b) Besitzen Musikstudierende an Ausbildungsstätten mit künstlerischem Schwerpunkt eine erhöhte Kompetenz und Leistungsmotivation im Vergleich zu Studierenden anderer Hochschulen und wenn Ja in welchen Bereichen?

3 Methode

3.1 Stichprobe

Insgesamt wurden 94 Studierende in musikbezogenen Fächern mit einem mittleren Alter von 23,4 Jahren ($SD=3,7$; $Md=22$) an folgenden Ausbildungsstätten erfasst: 1. Universität Gießen (Institut für Musikwissenschaften und Musikpädagogik): $n=59$ (16 Männer und 43 Frauen); Alter: $M=23,1$ ($SD=1,7$; $Md=22$); 2. Musikhochschulen Frankfurt und Stuttgart: $n=35$ (15 Männer und 20 Frauen); Alter: 24 ($SD=3,5$; $Md=22$).

3.2 Verfahren

Innerhalb von Seminaren und Vorlesungen wurden die Studierenden gebeten, freiwillig und unentgeltlich eine Fragebogenmappe auszufüllen und an einem zentralen Sammelort anonym abzugeben. Im Einzelnen wurden folgende Fragebögen verwendet:

- 1) Das Inventar zur Erfassung der Berufs- und Studiumsorientierung in der musikpädagogischen Ausbildung (IBSO), das grundlegenden Handlungs- und Motivationsfaktoren im Studium erfasst (vgl. von Georgi & Lothwesen, in Vorbereitung).

Der IBSO erfasst folgende Konstruktbereiche: *Empathie* (Soziale Orientierung und Einfühlungsvermögen in Studium und Beruf), *Stabilität* (Selbstsicherheit im Umgang mit dem Wissenserwerb und der Wissensanwendung und -produktion), *Motivation* (Musik als internaler Motivator für Studium und Beruf), *Kompetenz* (Theorie und Empirie als Kompetenzgrundlage für den späteren Beruf), *Resistenz* (Resistenz gegenüber sozialen und normativen Stressoren), *Flexibilität* (Allgemeine Offenheit für neue (musikalische) Impulse und Trends und deren praktische Anwendungen) und *Transfer* (die Bedeutung des Studiums (Referate, schriftliche Arbeiten) als Vorbereitung auf das spätere Berufsleben). Zunächst zeigte sich, dass die IBSO-Skalen miteinander unkorreliert sind ($p > 0,05$). Einzig die Skalen *Empathie* und *Motivation* ($r = 0,47$) sowie *Stabilität* und *Resistenz* ($r = 0,41$) korrelieren deutlich miteinander ($p < 0,01$). Eine nachträglich durchgeführte Clusteranalyse der IBSO-Skalen ergab, dass sich drei übergeordnete Cluster bilden lassen: a) Motivation (*Empathie* und *Motivation*), b) Stressbewältigungsressourcen (*Stabilität* und *Resistenz*) sowie c) fachbezogene Kompetenzen (*Kompetenz* und *Transfer*). Die Skala *Flexibilität* clustert nicht mit anderen Skalen zusammen (von Georgi & Lothwesen, in Vorbereitung).

- 2) Als Persönlichkeitsverfahren wurde die positiv und negativ Affekt-Skala (PANAS: Watson et al., 1988; deutsche Version: Krohne et al., 1996) (trait-Form) (vgl. Watson, 2000) sowie das Selbstkonzept Inventar (SKI: von Georgi & Beckmann, 2004) verwendet.
- 3) Um das habituelle Verhalten in Stresssituation zu erfassen wurde der Stressverarbeitungsfragebogen (SVF-S 44: Janke et al., 1995) vorgelegt.

3.3 Statistische Auswertungsmethoden

Die Zusammenhänge der IBSO-Skalen mit den Persönlichkeitsskalen (SKI, PANAS) und den Stressverarbeitungsstrategien (SVF-44) wurden mittels Korrelationsanalyse und multipler Regressionsanalyse untersucht. Die Frage nach den Abhängigkeiten zwischen Persönlichkeit, Stressverarbeitung, selbstbeschriebener Kompetenz und Leistungsmotivation (IBSO) mit der Ausbildungsform wurde mittels logistischer Regressionsanalysen überprüft (AV: Hochschulform; UV: Skalen der jeweiligen Fragebögen).

4 Ergebnisse

4.1 Persönlichkeitsvariablen und Stressverarbeitung

Insgesamt ergibt sich zunächst, dass im Falle der Skalen *Stabilität*, *Resistenz*, *Flexibilität* und *Transfer* Männer höhere Werte besitzen als Frauen, wobei diese Unterschiede im Falle der Skala *Resistenz* am deutlichsten hervortreten ($r=0,24$; $p\leq 0,01$). Im Einzelnen zeigt sich, dass die IBSO-Skala *Empathie* mit den Persönlichkeitsmerkmalen *Attraktivität* und zwischenmenschliches *Vertrauen* (Bindungsverhalten) des SKI kovariiert ($p\leq 0,01$). Im Rahmen der Stressbewältigung ergibt sich, dass Personen mit hohen Werten in der IBSO-Empathieskala zudem eine hohe positive Selbstinstruktion, geringe Resignation und eine erhöhte Tendenz zu Situationskontrollversuchen aufweisen (SVF) ($p\leq 0,05$). Besonders bedeutsam sind zudem ein geringes Maß an Fluchttendenzen und eine erhöhte Suche nach sozialer Unterstützung in Stresssituationen, die auch über die multiple Regression signifikant werden ($R=0,42$; $p\leq 0,001$). Diese Befunde unterstreichen sehr deutlich die starke soziale Orientierung, die durch die IBSO-Skala *Empathie* erfasst wird. Hohe Werte in der Skala *Stabilität* gehen hingegen vor allem mit einer erhöhten positiven Affektivität (PANAS) sowie mit Ich-Stärke, Attraktivität und Durchsetzung (SKI) einher, die alle Anteile des Persönlichkeitsmerkmals Extraversion erfassen ($p\leq 0,01$). Die Skala *Ich-Stärke* des SKI erfasst zudem einen negativen Anteil einer allgemeinen emotionalen Labilität und das Merkmal *Durchsetzung* des SKI Anteile von Psychotizismus (vgl. von Georgi, 2006). In Stresssituationen reagieren Personen mit hohen Werten in der Skala *Stabilität* vermehrt einer allgemeinen Bagatellisierung und einer positiven Selbstinstruktion (SVF) ($p\leq 0,001$). Besonders bedeutsam sind zudem die Tendenzen in Richtung Reaktions- und Situationskontrollversuche sowie eine geringe Resignationstendenz ($R=0,71$; $p\leq 0,001$). Diese Befunde deuten auf eine Verbindung von Aspekten der Situationskontrolle und positiver Orientierung bzgl. äußerlicher Stressoren im Falle hoher Werte in der IBSO-Skala *Stabilität* hin. Die signifikanten Betakoeffizienten bei gleichzeitig insignifikanter Korrelationen im Falle der Variablen *Vertrauen* und *Ordnungsliebe* des SKI deuten auf einen möglichen Supressoreffekt hin (vgl. Cohen et al., 2003). Die Skala musikbezogener *Motivation* korreliert nur sehr schwach mit den erhobenen Persönlichkeitsmerkmalen. Einzig eine Tendenz zur Bagatellisierung ist bei Personen mit hohen Motivationswerten deutlich zu verzeichnen ($p\leq 0,05$). Weiterhin zeigt sich, dass Personen mit einer hohen Resistenz gegenüber sozialen und normativen Stressoren (Skala *Resistenz*) nicht nur eine geringe Ängstlichkeit und

emotionale Labilität (NA, Ich-Stärke) aufweisen, sondern zudem auch extraversionäquivalente Persönlichkeitseigenschaften (PA, Attraktivität). Des Weiteren sind sie durch eine ganze Reihe von Bewältigungsstrategien gekennzeichnet, die wirksam gegen Belastungen eingesetzt werden. Von diesen sind vor allem zu nennen: Bagatellisierung, geringe Fluchttendenzen, geringes Grübeln (gedankliche Weiterbeschäftigung), positive Selbstinstruktion sowie eine sehr geringe Resignationstendenz ($p \leq 0,001$). Von besonderer Bedeutung ist bei Personen mit hohen Resistenzwerten die Fähigkeit, sich in Stresssituationen entspannen zu können ($R=0,31$; $p < 0,001$). Personen mit geringen Werten in der Skala *Resistenz* sind somit eher ängstlich und verfügen nur bedingt über erfolgversprechende Stressbewältigungsstrategien. Eine erhöhte Flexibilität und Offenheit gegenüber neuen musikalischen Impulsen geht einher mit geringer *Attraktivität* (man empfindet sich in sozialen Gruppen als eher uninteressant) und geringer *Durchsetzungsfähigkeit* (man legt sehr viel Wert auf soziale Kooperation) ($p \leq 0,01$). Dem entgegen stehen starke Flucht- und Resignationstendenzen sowie Reaktionskontrollversuche ($R=0,39$; $p=0,003$). Letztlich weist die Skala *Transfer* nur zwei Interkorrelationen auf, wovon die negative Beziehung zur Bewältigungsstrategie *Resignation* die wichtigere zu sein scheint ($R=0,28$; $p=0,001$). Personen die meinen, sie können besonders viel des Gelernten aus der Hochschule in die Praxis transferieren, reagieren in Stresssituation tatsächlich mit einer geringeren Resignationstendenz.

4.2 Der Zusammenhang mit der Hochschulform

Die Vorhersage der Hochschulzugehörigkeit (Universität vs. Musikhochschule) anhand der Persönlichkeitsskalen ergab ein signifikantes Nagelkerkes R-Quadrat von $NR^2=0,28$ mit $p[df=7; Chi^2=20,337]=0,005$. Insgesamt wurden 73,3% der Stichprobe der richtigen Hochschulform anhand der Persönlichkeitsskalen korrekt zugeordnet. Jedoch zeigt sich im Einzelnen, dass nur die PANAS-Skala *positive Affektivität* ($p=0,002$) und die SKI-Skala *Durchsetzung* ($p=0,01$) eine bedeutsame Vorhersage ermöglichen.

Die logistische Regressionsanalyse der Vorhersage der besuchten Hochschulform anhand der erhobenen Stressverarbeitungsstrategien (SVF) ergab ebenfalls ein signifikantes Ergebnis mit $NR^2=0,286$ ($p[df=11; Chi^2=20,985]=0,034$) (H-L-Test: $p[df=8; Chi^2=6,683]=0,571$). Allerdings liegt hier der prozentuale Anteil der richtig zugeordneten Personen bei nur 65,2 %. Die Inspektion der einzelnen Regressionskoeffizienten ergab entsprechend keinerlei Vorhersagebedeutsamkeit für eine der elf Coping-Skalen ($p > 0,10$). Die Stressverarbeitung spielt somit keine bedeutsame Rolle.

Ein deutlich signifikantes Ergebnis erbrachten hingegen die sieben IBSO-Skalen bei der Vorhersage der Hochschulform mit $NR^2=0,404$ ($p[df=7; Chi^2=28,907]<0,001$) und einer richtigen Zuordnung von 80,5%, Wie in Abbildung 1 zu erkennen ist, gibt es keine Unterschiede in den Skalen *Empathie* und *Motivation*. Deutliche und zum Teil hochsignifikante Unterschiede ergeben sich jedoch in den verbleibenden Skalen. So ist Abbildung 1 zu entnehmen, dass Studierende der Universität in ihrer Selbsteinschätzung eine höhere *Resistenz* und eine deutlich höhere *Flexibilität* sowie eine höhere *Kompetenz* ($p<0,01$) empfinden, Studierende an Musikhochschulen hingegen einen höheren *Transfer* und eine erhöhte *Stabilität* ($p<0,05$) angeben.

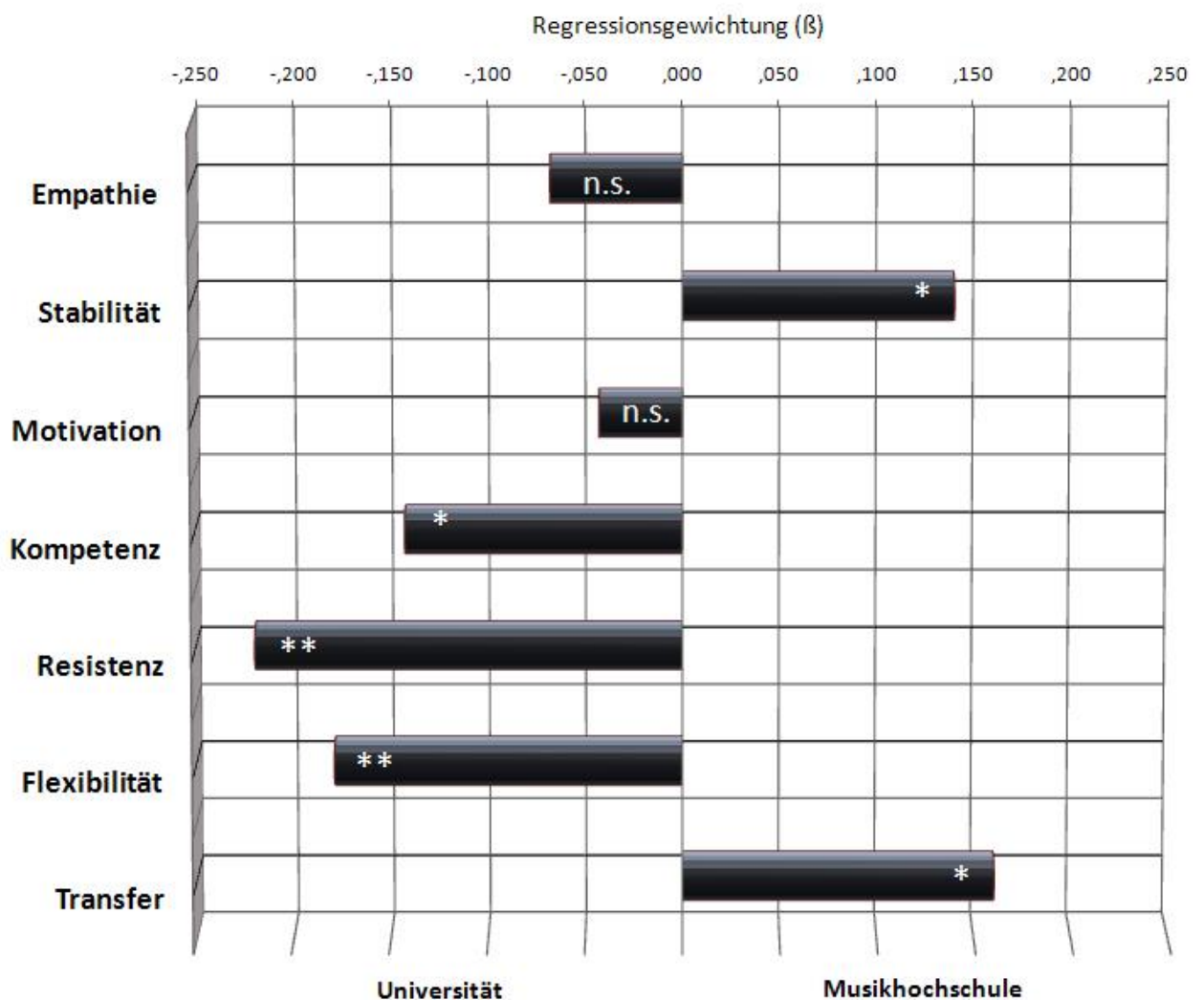


Abb. 1: Regressionsgewichtungen (β) für die Vorhersage der Zuordnung Universität (links) vs. Musikhochschule (rechts); (*): $p\leq 0,08$; (*): $p\leq 0,05$; (**): $p\leq 0,01$; (Regressionsstatistik: s. Text)

5 Diskussion

5.1 *Studiumsmotivation*

Die zwei IBSO-Skalen *Motivation* und *Empathie* können als grundlegende Handlungs- bzw. Studiumsmotivationsfaktoren interpretiert werden. Die selbsterlebte allgemeine soziale Orientierung spiegelt sich auch in den Korrelationen mit Persönlichkeitsvariablen wieder. So besitzen Personen mit entsprechenden hohen Werten ein großes soziales Netzwerk, ein deutlich positives Bindungsverhalten und ein positives Selbstwertgefühl. Ebenso zeigt sich, dass Stresssituationen nicht vermieden, sondern vor allem durch die Suche nach sozialer Unterstützung angegangen werden. Insgesamt spiegelt sich hier offenbar ein selbstempfundenes pro-soziales Verhalten und Empfinden, das für den Lehrerberuf von entscheidender Bedeutung ist und möglicherweise eine Grundmotivation für die Aufnahme des Studiums eines Musik vermittelnden Berufs darstellt. Dies haben in anderer Weise auch die Untersuchungen von Bastian (1995), Hoffmann (2007) und Neuhaus (2008) ermittelt. Aufgabe der Ausbildung ist entsprechend, dies zu fördern und ein Repertoire an sozialen Verhaltensoptionen in Lernsituationen zu eröffnen, um dies in der Berufspraxis umsetzen zu können (vgl. Niessen, 2006b).

Die zweite Dimension der Studiumsmotivation wird durch eine grundlegende Haltung gegenüber der allgemeinen Bedeutung von Musik im persönlichen Erleben gebildet. Personen, die Musik als grundlegenden Motivator betonen, neigen jedoch zur Bagatellisierung in Stresssituationen und weisen eine hohe Ich-Stärke auf. Fasst man dies zu einer möglichen Interpretation zusammen, so lässt sich vermuten, dass Personen mit sehr hohen Werten in dieser Dimension die Bedeutung von Musik überschätzen bzw. deren Bedeutung als interner Motivator nicht weiter differenzieren. So konnte die erziehungswissenschaftliche Forschung Zusammenhänge zwischen Studieninteresse und intrinsischer Motivation aufdecken (vgl. Lüders & Eisenacher, 2007: 138f). Folgte man dieser Interpretation so könnten sich bei Personen, die Musik sehr stark in den Mittelpunkt ihres motivationalen Erklärungsansatzes rücken, vielfältige persönliche Problemsituationen im beruflichen Alltag ergeben, in denen Musik nur noch eine geringe Rolle spielt und normative Zwänge (Lehrpläne, Unterrichtssituationen etc.) den Alltag bestimmen. Hans Günther Bastian problematisiert dies an studiengangsspezifischen Unterschieden bei der Einschätzung der Aussage „Musizieren ist mir eigentlich wichtiger als das Unterrich-

ten“ und warnt vor einem dadurch vorgezeichneten „Schulfrust“ (vgl. Bastian, 1995: 123).

5.2 Handlungskompetenzen der Stress- und Anforderungsbewältigung

Die Skalen *Stabilität* und *Resistenz* berühren vor allem jene Bereiche, bei denen es sich um Situationen mit einem Anforderungscharakter handelt. Hierbei lassen sich zwei Formen unterscheiden: 1. Anforderungen, die mit dem Wissenserwerb und dessen Präsentation (Referate, Hausarbeiten, Prüfungen) in Zusammenhang stehen (*Stabilität*) und mit allgemeiner Selbstsicherheit einhergehen und 2. Anforderungen und Stressoren, die aus sozialen und normativen Bedingungen des Studiums entstehen (Kritik von Mitstudierenden, Leistungsbewertung, Überlastung) (*Resistenz*). Obwohl beide Skalen miteinander positiv korrelieren liegt der qualitative Unterschied vor allem darin, dass die Skala *Stabilität* zuvorderst Ich-bezogene Anforderungen aufgreift, während die Skala *Resistenz* auf soziale und normative Bewertungsprozesse abzielt. Eine solche Trennung erscheint durchaus sinnvoll und auch für weitere Forschungsarbeiten und Überlegungen sehr gewinnbringend. Die Interkorrelationen zeigen, dass Personen dann ein positives Selbstbild aufweisen, wenn sie in ihrer Selbsteinschätzung eine gleichermaßen hohe *Stabilität* und *Resistenz* besitzen. Allerdings zeigt sich auch, dass Studierende, die eine hohe *Stabilität* empfinden, zugleich eine allgemeine positive Affektivität (PA) aufweisen und vor allem Copingstrategien verwenden, die eine selbsterlebte Kontrolle herstellen. Studierende, die eine hohe *Resistenz* besitzen, sind gekennzeichnet durch das Fehlen einer negativen Affektivität (NA) bzw. einer emotionalen Labilität und verwenden Stressbewältigungsstrategien, die eher außerhalb der direkten Situationsbeeinflussung liegen und interne kognitive Prozesse betreffen (Bagatellisierung, Entspannung, kein Grübeln). Interessanterweise sehen die von Hoffmann befragten Studierenden den Schwerpunkt ihrer Qualifikation für den Musiklehrerberuf vorrangig im „sozial-kommunikativen, menschlichen, emotionalen Bereich“, was Hoffmann zusammenfassend bilanziert: „Es wird deutlich, dass die Befragten sich letztendlich auf dieses Potenzial verlassen und nicht auf fachliche Fähigkeiten, wenn es darum geht, auch perspektivisch ihre Eignung für den musikpädagogischen Beruf zu belegen.“ (Hoffmann, 2007: 73). Die hier unterschwellig artikulierte Gefahr der Überbetonung subjektiver personaler Qualitäten ist im Zusammenhang mit möglichen Belastungsfaktoren im Berufsalltag zu diskutieren, die ein Ungleichgewicht zwischen inneren Ansprüchen und äußeren Anforderungen bedingen können. So beschreibt Bastian als idealistisches Hauptmotiv der Studienwahl ein „emotio-

nal befriedigendes Arbeiten in der Schule“, warnt jedoch gleichzeitig vor einem möglichen Burn-out-Risiko im Falle einer Nicht-Erfüllung dieses selbstformulierten Anspruchs (Bastian, 1995: 118).

Möglicherweise kovariieren beide Dimensionen nicht nur mit den gegebenen kognitiven Fähigkeiten, sondern auch und insbesondere mit dem seitens der Ausbildungsstätten gestellten Anforderungsniveau. Die Gefahr, dass ein subjektiv gering empfundenes Anforderungsniveau ein hohes subjektives Leistungsniveau suggeriert, kann erhebliche Probleme im Beruf nach sich ziehen. Verstärkt werden kann dieses erhöhte Burn-out-Risiko möglicherweise durch vorliegende soziale Bewertungsängste (vgl. Rauin & Maier, 2007). Dies verweist auf einen leider viel zu selten diskutierten Sachverhalt: Dort wo Studierende scheinbar ‘von selbst’ mögliche praktische Sachverhalte und Kompetenzen erlernen, ohne dass Anforderungen bestehen, die spürbar oberhalb des Leistungsniveaus liegen (vgl. Expertiseforschung (Ericsson et al., 2009)), besteht die Gefahr der Etablierung falscher Vorstellungen über das eigene tatsächliche Leistungs- und Kompetenzniveau.

5.3 *Fachbezogene Kompetenzen*

Die Frage nach der erlebten fachbezogenen Kompetenz, wird durch die IBSO-Skalen *Kompetenz* und *Transfer* erfasst. Die Kompetenzdimension des IBSO beinhaltet vor allem eine grundlagenwissenschaftliche und empirisch ausgerichtete Grundhaltung der Studierenden, die Transferdimension eine subjektive Einschätzung der Wichtigkeit und Übertragbarkeit des Gelernten in die spätere berufliche Praxis.

Die Interkorrelationen der Persönlichkeitsskalen mit der IBSO-Skala *Kompetenz* spiegeln zunächst eine interessante Tatsache wieder: Eine hohe subjektive grundlagenwissenschaftliche und empirisch-statistische Orientierung geht einher mit einer leichten Zwanghaftigkeit (Ordnungsliebe) und vor allem mit Reaktionskontrollversuchung als Stressverarbeitungsstrategie. Auch hier kann ein Bezug zur Expertisierung hergestellt werden, indem eine zwanghafte und anstrengende Fokussierung auf wissenschaftlich-analytisches Arbeiten, was zugleich anstrengend und unangenehm sein kann, häufig vermieden wird. In diesem Zusammenhang ist die von Neuhaus faktorenanalytisch gebildete Dimension „wissenschaftliches Arbeiten“ interessant, die offensichtlich von deutlich nachrangiger Bedeutung bei der Bewertung der Studieninhalte ist (vgl. Neuhaus, 2008: 232). Ähnlich kann der Befund von Bastian interpretiert werden, der generell höchste Zustimmungen der Studierenden bei Schulprak-

tika nachweisen konnte, wobei Fachwissenschaft, psychologische und soziologische Bereiche, Musiktheorie und Analyse sowie Gehörbildung nur mittlere Rangplätze einnahmen (vgl. Bastian, 1995: 104).

Die Skala *Transfer* erfasst, in wie weit die befragten Personen glauben, dass das im Studium Gelernte eine Grundlage für die spätere Praxis darstellt. Hierin äußern sich Vorstellungen von der späteren Berufspraxis, die bei den Studierenden aufgrund unterschiedlicher Fachsemester und damit einhergehenden schulpraktischen Erfahrungen oftmals noch nicht ausgeprägt sein können. In diesem Sinne können sich die Befunde der vorliegenden Studie einfügen zwischen retrospektive Befragungen von bereits berufserfahrenen Musiklehrkräften (vgl. Niessen, 2006a) und Erhebungen der Berufsvorstellungen von Studierenden (Bastian, 1995; Neuhaus, 2008). Eine direkte Vergleichbarkeit der Ergebnisse ist jedoch nur mit Einschränkung möglich, da die unterschiedlichen Perspektiven dieser Ansätze Verzerrungen erwarten lassen. Dennoch lassen sich aus den Befunden einige interessante Hypothesen und Forschungsfragen ableiten. Die hier erfasste Tendenz der Übertragung von erlernten Studieninhalten in die Berufspraxis (*Transfer*) ist verbunden mit einer geringeren Resignationstendenz und mit Reaktionskontrollversuchen, und sie erscheint weitestgehend unabhängig von der Persönlichkeit. Mit Blick auf zukünftige Forschungen ist anzunehmen, dass eine subjektiv erlebte Praxisrelevanz der im Studium erarbeiteten Kompetenzen für Vorträge, schriftliche Arbeiten etc. zu einer erhöhten Motivation und positiven Ausrichtung zu Beginn der beruflichen Tätigkeit führt. In wie weit eine derartige subjektiv erlebte Orientierung innerhalb des Studiums sich tatsächlich als hilfreich und motivierend auswirkt oder aber langfristig eher zum Gegenteil führt, ist in der vorliegenden Studie nicht zu klären. Ergebnisse der Langzeitstudie von Rauin und Maier zur retrospektiven Einschätzung unterschiedlicher Lernbereiche zeigen tendenziell abweichende Kompetenzprofile und -erfahrungen in Studium und zweiter Ausbildungsphase bei den befragten Lehramtsstudierenden und deren individuelles Engagement (Rauin & Maier, 2007: 116ff).

5.4 *Flexibilität und Offenheit*

Die IBSO-Skala *Flexibilität*, die einen offenen Umgang mit neuen Dingen und Themen erfasst sowie die Fähigkeit sich auf neue Situationen einzustellen und mit neuen Materialien den Unterricht zu gestalten, scheint eine wichtige Sonderrolle innerhalb der Motivations- und Handlungskompetenzen einzunehmen, die in den meisten Studien bisher nicht berücksichtigt wurde. Möglicherweise weist diese Skala einen engen Bezug zu äquivalenten Persönlichkeitsdimensi-

onen auf (z.B. *Offenheit für neue Erfahrungen* des NEO-ffi (Costa & McCrae, 1985), *experience seeking* (Zuckerman, 1994) oder Offenohrigkeit (Gembris, 2002; von Georgi, 2010). Derartige Konstrukte sind gekennzeichnet durch die Tendenz, neuen Erfahrungen und Erlebnissen gegenüber aufgeschlossen zu sein und aktiv zu suchen. Interessanterweise ist die Skala *Flexibilität* verbunden mit einer geringen Fähigkeit, eigene Interessen gegen Andere durchzusetzen und geht einher mit dem Gefühl der Unattraktivität (Marginalität) in sozialen Gruppen. Im Stressverarbeitungsbereich besteht zudem eine erhöhte Tendenz zur Resignation. Das bedeutet, dass zwar die Kompetenz und auch das generelle Interesse zur Hinwendung auf neue oder kritische Themengebiete besteht, die praktische Umsetzung derartiger Themen aber bei Auftreten von sozialen Konflikten (z.B. mit Kollegen oder Schülern) durch deutliche Resignations- und Rückzugstendenzen gefährdet sein kann. Die Dimension *Flexibilität* ist daher möglicherweise von erheblicher Praxisrelevanz, da sie den Umgang mit Neuem und die Kommunikation darüber innerhalb bestehender sozialer Systeme hervorhebt. Der konsequenten praktischen Umsetzung von Ideen stehen hierbei Resignationstendenzen einerseits und eine erhöhte Kooperationsbereitschaft andererseits gegenüber. Letzteres erscheint im Studium durchaus sinnvoll, birgt jedoch im Berufsalltag die Gefahr des Scheiterns (vgl. Blutner, 2004: 150f.). Vor diesem Hintergrund erscheint es sinnvoll im Rahmen der Ausbildung Projektseminare anzubieten, die vor allem darauf abzielen, Kompetenz und Selbstsicherheit bei der Präsentation und Implementierung neuer oder alternativer Konzepte zu erlernen (vgl. Niessen, 2006b).

Personen mit geringen Werten in der Dimension *Flexibilität* erscheinen eher durchsetzungsfähig, zwanghaft, empfinden sich in sozialen Gruppen als attraktiv und weisen keine Stressvermeidungsstrategien auf. Diese Personen dürften vor allem Schwierigkeiten bekommen, wenn es darum geht, neue musikalische Strömungen zu erkennen und deren Bedeutung zu vermitteln. Auch der Erwerb und die Umsetzung neuer curricularer Ansätze dürften für diese Personen schwierig sein – wenn sie überhaupt angegangen werden. Sicherlich kann eine solche Orientierung durchaus einem Burn-out und weiteren vielfältigen Konfliktsituationen in der späteren beruflichen Praxis entgegenwirken – offen bleibt hierbei jedoch die pädagogische Eignung derartiger Personen (Rauin und Maier sprechen in einem ähnlichen Zusammenhang von „Ressourcen schonenden Routinen“ (Rauin & Maier, 2007: 116)).

5.5 Unterschiede zwischen Studierenden an Universität und Musikhochschulen

Zwischen den Studierenden an Universität und Musikhochschulen ergeben sich zunächst keine Unterschiede in jenen Bereichen, die Anteile der Studiumsmotivation erfassen (*Empathie* und *Motivation*). Somit ist die innersubjektive Orientierung in Richtung eines musikbezogenen Berufes, sowohl was das soziale Einfühlungsvermögen als auch die musikalische Grundmotivation anbelangt, in beiden Gruppen gleich stark vorhanden und nicht abhängig vom Ausbildungsort oder von der Ausrichtung der jeweiligen Fächer. Dies hat auch Neuhaus in ihrer Studie ermittelt: hinsichtlich der pädagogischen und künstlerischen Orientierung war kein Unterschied zwischen Studierenden an Musikhochschulen und Universitäten festzustellen (Neuhaus, 2008: 233). *Empathie* und *Motivation* können somit als Basisvariablen für ein musikbezogenes Fach gesehen werden, die letztlich Erfolg und Misserfolg in Studium und Beruf grundsätzlich determinieren. So bestimmt die musikbezogene Motivation einerseits die Fächerwahl, geht jedoch andererseits einher mit einer allgemeinen sozialen Orientierung, die nicht nur wichtig für das praktische Musizieren ist, sondern auch für die Entwicklung der berufspraktischen Kompetenz, Musik in Gruppenkontexten zu vermitteln. Hinzu kommt, dass eine soziale Orientierung innerhalb der Ausbildung für viele eine wichtige Voraussetzung ist, die Herausforderung eines Studiums zu bestehen (vgl. hierzu den Studierendentypus der „geselligen, pragmatischen Hedonisten“, Rauin & Maier, 2007: 125). Im Gegensatz zu bestehenden Untersuchungen, verweist dieser Befund auf einen erweiterten Blickwinkel zur Erklärung und Vorhersage des Verhaltens und Erlebens im späteren Lehrerberuf: Während die Aufgabe der musikbezogenen Motivation mit hoher Wahrscheinlichkeit letztlich zu einem Berufswechsel führen wird, scheint der sozialen Orientierung eine wichtige Rolle für die berufliche Zukunft zuzukommen. Somit ließe sich die Hypothese aufstellen, dass nicht etwa die Studiumsmotivation an sich von entscheidender Bedeutung ist, sondern vielmehr eine grundsätzliche soziale Orientierung, die im Rahmen des Studiums erlebt und weiter entwickelt werden muss. Gelingt dieses nicht, so kann es später bei Belastungen im Berufsalltag sehr schnell zum Burn-out kommen. Auch aus stresstheoretischer Sicht ist dieser Ansatz zu untermauern (Cohen & Wills, 1985). In diesem Zusammenhang kann diskutiert werden, ob die immer stärkere Leistungsausrichtung und Überprüfung sowie Fachreglementierung („Verschulung“) an den Hochschulen, im Rahmen des Bologna-Prozesses, einer entsprechenden sozialen Kompetenz- und Persönlichkeitsentwicklung nicht entgegensteht – sofern man nicht die Ansicht vertritt, dass be-

berufliche Stresssituationen einzig durch Fachkompetenz vermieden oder bewältigt werden könnten.

Deutliche Unterschiede zwischen den Ausbildungsstätten Universität und Musikhochschule ergaben sich in den verbleibenden IBSO-Skalen. Zunächst scheinen sich beide Ausbildungsstätten in der subjektiven Einschätzung bestehender Handlungs- und Studiumsmotivationsfaktoren dahingehend zu unterscheiden, dass Studierende an Musikhochschulen sich eher in der Lage sehen, das im Studium geforderte Wissen zu etablieren und zu präsentieren. Zudem äußern sie ein scheinbar höheres Transferempfinden bzgl. der Übertragbarkeit des im Studium Erlernten in die berufliche Praxis. An dieser Stelle sei nochmals auf parallele Tendenzen der Ergebnisse aus der Studie von Hoffmann (2007) verwiesen und die von ihr festgestellten höheren Werte der Schulmusikstudierenden bzgl. der Selbstwirksamkeit; auch in dieser Studie wurden jedoch Unterschiede zwischen den Ausbildungsstätten nicht explizit untersucht. Studierende an Universitäten sind hingegen dadurch gekennzeichnet, dass sie sich in den genannten Bereichen eher geringere selbstbeschriebene Kompetenzen zuschreiben, gleichzeitig jedoch offener für neue Ansätze sind und zudem eine deutlich geringere Angst vor sozialen Bewertungsprozessen aufweisen sowie eine erhöhte empirisch-wissenschaftliche Kompetenz. In wie weit diese Unterschiede das Resultat des entsprechenden Lehr-Lernniveaus einer Ausbildungsstätte darstellen, ist im Augenblick nicht abzuschätzen. Weiterhin war eine verringerte Flexibilität vor allem bei Studierenden der Musikhochschulen zu beobachten, was durchaus kritisch zu sehen ist, da diese gleichsam eine erhöhte Praxisorientierung bzw. Kompetenzen zu besitzen äußerten. In dieser Hinsicht wäre das jeweilige Lehrangebot auf musikstilistische und didaktische Vielfalt und Offenheit zu prüfen.

6 Ausblick

Zusammenfassend kann festgestellt werden, dass die Versuchspersonen der Universität durch eine eher wissenschaftlich orientierte Handlungskompetenz gekennzeichnet sind, Studierende an Musikhochschulen eher durch eine praxisbezogene Handlungskompetenz. Ohne an dieser Stelle auf die unterschiedlichen Ausbildungsformen eingehen zu wollen, zeigt sich somit sehr deutlich, dass sich, trotz zum Teil identischer Lehr- und Lerninhalte, die subjektiven Wahrnehmungen der befragten Studierenden erheblich unterscheiden. Hierbei zeigen die Ergebnisse, dass die Motivations- und Handlungskompetenzen durch eine Reihe von Variablen und Ausbildungsfaktoren beeinflusst werden,

deren Auswirkungen in nachfolgenden Untersuchungen weiter überprüft werden müssen. So erscheinen die komplexen Interaktionen zwischen einer persönlichkeitsabhängigen Kompetenz- und Motivationsentwicklung, Stressverarbeitungsstrategien und dem Anspruchsniveau der Ausbildungsstelle von wichtiger Bedeutung für das (spätere) subjektive Erleben im Beruf zu sein. Und somit muss festgestellt werden, dass eine alleinige Modifikation objektiver Bedingungen an musikbezogenen Ausbildungsstätten (Anforderungen, Module, Prüfungsinhalte etc.) nur mit Einschränkung in der Lage sind automatisch eine höhere Kompetenz in unterschiedlichen Bereichen herbeizuführen.

Danksagung

Für ihre Mitarbeit an der Erstellung der Fragebögen und Datenerhebung danken wir Tabea Köhler, Matthias Schmitt und Marit Stahl.

Literatur

- Bandura, A. (1997). *Self-Efficacy – The Exercise of Control*. New York: Freeman and Company.
- Bastian, H. G. & Kraemer, R.-D. (1992). *Musikpädagogische Forschung in Deutschland. Dokumentation und Analyse*. Mainz: Schott.
- Bastian, H. G. (1995). Studien- und Berufsmotivation von Musiklehrerstudentinnen und -studenten. Ergebnisse einer Befragung in den alten und neuen Bundesländern. In N. Knolle & Th. Ott (Hrsg.). *Zur Professionalisierung von Musiklehrern. Ausbildungsprobleme in den alten und neuen Bundesländern*. (S. 91–154) Mainz: Schott.
- Behne, K.-E. (1979). Motivation bei Musikstudenten. Bericht über einen Forschungsprozeß. *Motivationsforschung in der Musikpädagogik* (=Forschung in der Musikpädagogik 1979, S. 92–97). Mainz: Schott.
- Blutner, D. (2004). Führungskompetenz im Mitgliedschaftsdilemma. In W. Böttcher & E. Terhart (Hrsg.). *Organisationstheorie in pädagogischen Feldern. Analyse und Gestaltung* (S. 142–158). Wiesbaden: Verlag für Sozialwissenschaften.
- Bouij, Ch. (2004). Two Theoretical Perspectives on the Socialization of Music Teachers. *Action, Criticism, and Theory for Music Education*, 3 (3), Dec

2004

URL: http://act.maydaygroup.org/articles/Bouij3_3.pdf (Abruf 4.3.2010)

- Cohen, J., Cohen, P., West, S. G. & Aiken, L. S. (2003). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences* (3. Aufl.). Mahawah, New Jersey: Lawrence Erlbaum Associates.
- Cohen S & Wills T A. (1985). Stress, social support, and the buffering hypothesis. *Psychological Bulletin*, 98, 310-57.
- Costa, P. T. & McCrae, R. R. (1985): *The NEO- Personality Inventory Revised Manual (NEO-PI-R)*. Odessa, Florida: Psychological Assessment Resources 1985. [Deutsche Fassung: Borkenau, P. & Ostendorf, F. (1993): *NEO-Fünf-Faktoren Inventar (NEO-FFI)*. Göttingen: Hogrefe.
- Costello, A. B. & Osborne, J. W. (2005). Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most From Your Analysis. *Practical Assessment, Research & Evaluation*, 10 (7), URL: <http://pareonline.net/pdf/v10n7.pdf> (Abruf 10.5.2010).
- Darius, P. (1976). Die Gliederung des Lehrangebots in der Musikerziehung an Pädagogischen Hochschulen und die Studienerwartungen der Studenten. Eine empirische Untersuchung an sechs Hochschulen in Nordrhein-Westfalen. *Musikpädagogik in der Studienreform* (=Forschung in der Musikpädagogik 1976, S. 347–353). Mainz: Schott.
- Ericsson, K. A., Nandagopal, K & Roring, R. W. (2009). Toward a science of exceptional achievement. *Longevity, Regeneration, and Optimal Health: Annals of the New York Academic Science*, 1172, 199-217.
- Gembris, H. (2002). *Grundlagen musikalischer Begabung und Entwicklung*. Augsburg: Wießner.
- Grimmer, F. (1986). Künstlerische Ausbildung und Lebensgeschichte. Zur Konstitution persönlich bedeutsamer Lernprozesse im Klavierunterricht bei Musikstudierenden in der Lehrerbildung. In H. J. Kaiser (Hrsg.): *Unterrichtsforschung* (=Musikpädagogische Forschung, Bd. 7, S. 71–86) Laaber: Laaber.
- Grimmer, F. (1987). Klavierausbildung im Spiegel subjektiver Deutung. Zur Auseinandersetzung mit eigener Lern- und Bildungsgeschichte von Musikstudierenden in der Lehrerbildung. In: G. Kleinen (Hrsg.). *Außer-schulische Musikerziehung* (=Musikpädagogische Forschung, Bd. 8, S. 65–77). Laaber: Laaber.
- Gulliksen, H. (1950). *Theory of Mental Tests*. New York: Wiley.

- Hilligus, A. H. & Rinkens, H.-D. (Hrsg.) (2006). *Standards und Kompetenzen - neue Qualität in der Lehrerausbildung? Neue Ansätze und Erfahrungen in nationaler und internationaler Perspektive*. Münster: Lit.
- Hoffmann, G. (2007). Qualifikationsmerkmale von Studierenden der Musikpädagogik in Selbstzeugnissen. In G. Hoffmann (Hrsg.). *Identität und Kreativität. Beiträge aus musikwissenschaftlicher und musikpädagogischer Forschung* (S. 55–78). Augsburg: Wißner.
- Hörmann, K. (1981). Fragen zur musikalischen Sozialisation von Studienanfängern. In: K.-E. Behne (Hrsg.). *Musikalische Sozialisation* (=Musikpädagogische Forschung, Bd. 2, S. 104–133) Laaber: Laaber.
- Janke, W., Erdmann, G., Kallus, K. W. & Boucsein, W. (1995). *Stressverarbeitungsfragebogen – SVF-S 44*. Göttingen: Hogrefe.
- Kadushin, Ch. (1969). The Professional Self-Concept of Music Students. *The American Journal of Sociology*, 75 (3), 389–404.
- Kraemer, R.-D. (Hrsg.)(1991). *Musiklehrer. Beruf, Berufsbild, Berufsverlauf*. (=Musikpädagogische Forschung, Bd. 12) Essen: Die Blaue Eule.
- Krohne, H. W., Egloff, B., Kohlmann, C.-W. & Tausch, A. (1996). Untersuchungen mit einer deutschen Form der Positive und Negative Affect Schedule (PANAS). *Diagnostica*, 42, 139–156.
- Lord, F. M. & Novick, M. R. (1968). *Statistical Theories of Mental Test Scores*. Reading - Massachusetts, Menlo Park - California, London, Don Mills - Ontario: Addison-Wesley Publishing.
- Lüders, M. & Eisenacher, S. (2007): Zeitlicher Studieraufwand von Studierenden. In M. Lüders & J. Wissinger (Hrsg.). *Forschung zur Lehrerbildung. Kompetenzentwicklung und Programmevaluation*. (S. 133-150) Münster: Waxmann.
- Lüders, M. & Wissinger, J. (Hrsg.)(2007). *Forschung zur Lehrerbildung. Kompetenzentwicklung und Programmevaluation*. Münster: Waxmann.
- Müller, R. (1991). Erfolgstyp Musiklehrer. Dimensionen der Interaktionskompetenz. In R.-D. Kraemer (Hrsg.). *Musiklehrer. Beruf, Berufsbild, Berufsverlauf*. (=Musikpädagogische Forschung, Bd. 12, S. 90–100) Essen: Die Blaue Eule.
- Neuhaus, D. (2007). „Zur Zeit sehe ich meine berufliche Zukunft als...“ Ergebnisse einer Befragung zum Berufswahlprozess von Lehramtsstudierenden mit dem Fach Musik. In N. Schläbitz (Hrsg.). *Interkulturalität als*

Gegenstand der Musikpädagogik (=Musikpädagogische Forschung, Bd. 28, S. 287–304) Essen: Die Blaue Eule.

- Neuhaus, D. (2008). *Perspektive Musiklehrer/in. Der Berufswahlprozess von Lehramtsstudierenden mit dem Unterrichtsfach Musik*. Köln: Dohr.
- Nielsen, F. V. (2004). Music Teacher Education in a Field of Tension between Science and Art and between Praxis and Theory. An Introduction to a Discussion on Music Teacher Education as Professionalization. In F. Niermann & C. Wimmer (Hrsg.). *Musiklernen - ein Leben lang*. (S. 195-200) Wien: Universal Edition.
- Niessen, A. (2006a). *Individualkonzepte von Musiklehrern*. Münster: Lit
- Niessen, A. (2006b). Musiklehrer werden, sein und bleiben – Profession und Persönlichkeit in Ausbildung und Berufsalltag. *Diskussion Musikpädagogik*, 32 (6), 4–14.
- Oser, F. & Oelker, J. (2001). Standards: Kompetenzen von Lehrpersonen. In F. Oser & J. Oelkers (Hrsg.). *Die Wirksamkeit der Lehrerbildungssysteme. Von der Allrounderbildung zur Ausbildung professioneller Standards*. (S. 215–342) Chur u.a.: Rüegger.
- Rauin, U. & Maier, U. (2007). Subjektive Einschätzungen des Kompetenzerwerbs in der Lehramtsausbildung. In M. Lüders & J. Wissinger (Hrsg.). *Forschung zur Lehrerbildung. Kompetenzentwicklung und Programmevaluation*. (S. 103-132) Münster: Waxmann.
- Roulston, K., Legette, R. & Womack, S. T. (2005). Beginning Music Teachers' Perceptions of the Transition from University to Teaching in Schools. *Music Education Research*, 7 (1), 59–82.
- Schulten, M. L. (1979). *Das Berufsbild des Musiklehrers*. Hamburg: Wagner.
- Schulten, M. L. (2004). Musiklehrerausbildung zwischen Professionalität der Musikpädagogik als Wissenschaft und professioneller Praxis. In F. Niermann & C. Wimmer (Hrsg.). *Musiklernen - ein Leben lang*. (S. 219–222) Wien: Universal Edition.
- Siedenburg, I. (2009). *Geschlechtstypisches Musiklernen. Eine empirische Untersuchung zur musikalischen Sozialisation von Studierenden des Lehramts Musik*. Osnabrück: epOs.
- Stroh, W. M. (1996). Geschlechtsstereotype Tendenzen in chaotischen Systemen: Frauen und Männer im Oldenburger Musikstudium. In H. J. Kaiser

- (Hrsg.). *Geschlechtsspezifische Aspekte des Musiklernens* (=Musikpädagogische Forschung, Bd. 17, S. 110–122) Essen: Die Blaue Eule.
- Terhart, E. (2007). Erfassung und Beurteilung der beruflichen Kompetenz von Lehrkräften. In M. Lüders & J. Wissinger (Hrsg.). *Forschung zur Lehrerbildung. Kompetenzentwicklung und Programmevaluation*. (S. 37–62) Münster: Waxmann.
- Vogt, J. (2002). „Praxisbezug als Problem. Zur Professionalisierung der Musiklehrerausbildung“. *Zeitschrift für kritische Musikpädagogik*, August 2002. URL: <http://home.arcor.de/zf/zfkm/vogt4.pdf> (Abruf 4.3.2010)
- Vogt, J. (2004). What Does ‘Professionalization’ Mean in Music Teacher Education? In F. Niermann & C. Wimmer (Hrsg.). *Musiklernen - ein Leben lang*. (S. 201–209) Wien: Universal Edition.
- von Georgi, R. (2006). *Theorie und Messung subjektiver Beschwerden*. Osnabrück: Der Andere Verlag.
- von Georgi, R. (2010). Persönlichkeitstheorien. In H. de la Motte-Haber, H. von Loesch, G. Rötter & C. Utz (Hrsg.). *Lexikon der Systematischen Musikwissenschaft*. (S. 373–377). Laaber: Laaber Verlag.
- von Georgi, R. & Beckmann, D. (2004). *SKI: Selbstkonzept Inventar*. Bern: Huber.
- von Georgi, R. & Lothwesen, K. S. (in Vorbereitung). Konstruktion eines Tests zur Erfassung der Handlungskompetenzen und Studiumsmotivation in der musikpädagogischen Ausbildung.
- Watson, D.; Clark, L. A. & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54, 1063–1070.
- Watson, D. (2000). *Mood and temperament*. New York, London: The Guilford Press.
- Zimmerschied, D. (1978). Feststellungen zum Bedarf an Musiklehrern und Untersuchungen über die Motivation zur Studienwahl. In H. W. Höhnen, B. Binkowski, H. Hopf & R. Jakoby (Hrsg.). *Entwicklung neuer Ausbildungsgänge für Lehrer der Sekundarstufen I und II*. (S. 15–58) Regensburg: Bosse.
- Zuckerman, M. (1994). *Behavior expressions and biosocial bases of sensation seeking*. New York: Cambridge University Press.